Developing Retail Performance Measurement and Financial Distress Prediction Systems by Using Credit Scoring Techniques

Yu-Chiang Hu

Doctor of Philosophy

The University of Edinburgh

2006

Declaration

This thesis is composed by me and that the work is my own. No part of it has been submitted to any other institution for another qualification.

Signature:

Date: 6th Sept 2006

Abstract of Thesis

The current research develops a theoretical framework based on the Resource-Advantage Theory of Competition (Hunt, 2000) for the selection of appropriate variables. Using a review of the literature as well as to interviews and a survey, 170 potential retail performance variables were identified as possible for inclusion in the model. To produce a relative simple model with the aim of avoiding over-fitting, a limited number of key variables or principal components were selected to predict default. Five credit-scoring techniques: *Naïve Bayes*, *Logistic Regression*, *Recursive Partitioning*, *Artificial Neural Network*, and *Sequential Minimal Optimization (SMO)* were employed on a sample of 195 healthy and 51 distressed businesses from the USA market over five time periods: 1994-1998, 1995-1999, 1996-2000, 1997-2001 and 1998-2002.

Analyses provide sufficient evidence that the five credit scoring methodologies have sound classification ability in the year before financial distress. Moreover, they still remained sound even five years prior to financial distress. However, it is difficult to conclude which modelling technique has the highest classification ability uniformly, since model performance varied in terms of different time scales. The analysis also showed that external environment influences do impact on default assessment for all five credit-scoring techniques, but these influences are weak. These findings indicate that the developed models are theoretically sound. There is however a need to compare their performance to other approaches.

To explore the issue of the model's performance two approaches are taken. First, rankings from the study were compared with those from a standard rating system—in this case the well-established Moody's Credit Rating. It is assumed that the higher the degree of similarity between the two sets of rankings, the greater the credibility of the prediction model. The results indicated that the logistic regression model and the SMO model were most comparable with Moody's. Secondly, the model's performance was assessed by applying it to different geographical areas. The original USA model was therefore applied to a new US data set as well as the European and Japanese markets. Results indicated that all market models displayed similar discriminating ability one year prior to financial distress. However, the USA model performed relatively better than European and Japanese models five years before financial distress. This implied that a financial distress model has potentially better prediction ability when based on a single market.

Following this result it was decided to explore the performance of a generic global model, since model construction is time-consuming and costly. A composite model was constructed by combining data from USA, European and Japanese markets. This composite model had sound prediction performance, even up to five years before financial distress, as the accuracy rate was above 85.15% and AUROC value was above 0.7202. Comparing with the original USA model, the composite model has similar prediction performance in terms of the accuracy rate. However, the composite model presented a worse prediction utility based on the AUROC value. A future research direction might be to include more world retailing markets in order to ensure the model's prediction utility and practical applicability.

Acknowledgements

This thesis is funded by the College of Humanities and Social Science, the Management School at the University of Edinburgh as well as the Overseas Research Students Awards Scheme (ORSAS). I am very grateful to have the opportunity to do my research in such an excellent learning environment, and would like to express my sincere gratitude to a number of people, who have supported me to finish this thesis.

First and foremost, I would like to thank my research supervisor Professor Jake Ansell. Whenever I had any questions relating to my research, he was always available and gave me invaluable suggestions and inputs. Moreover, he has always inspired me to develop innovative thinking and encouraged me to face the challenges. His encouragement has given me confidence in what I do. Professor Jake Ansell's guidance has allowed me to complete my thesis in a rigorous manner. My deepest thanks to Professor Jake Ansell for all his continued support through the three years.

Second, I wish to express my appreciation to Professor John Dawson for his kind assistance towards my research questionnaire design and data collection. I would also like to thank all faculty members at the management school who have helped me along the way.

Third, I am very grateful to all interview and survey participants in this research. They provided me much valuable information to enhance my research findings. Many thanks also to my PhD colleagues and friends for making this PhD an enjoyable journey.

Finally, I am most thankful to my family. A special note of gratitude to my wife Fatima Wang for her patients and love, and to my mother and brothers for their unique moral support.

Table of Contents

Part One: Background of Thesis

Chapter ONE
Introduction

Part Two: Review of the Literature

Chapter TWO
Development of Performance Measurement and Prediction System

Part Three: Research Framework Development

Chapter THREE
Resource-Advantage Theory of Competition

Chapter FOUR
Research Framework Development

Chapter FIVE
Survey Examination of the Research Framework

Part Four: Default Prediction Model Construction

Chapter SIX
Model Construction: Data Collection and Key Variables Determination

Chapter SEVEN
Model Construction: Modelling Techniques and Cross-Validation Process

Part Five: Model Utility Evaluation

Chapter EIGHT
Model Prediction Performance Evaluation

Chapter ELEVEN

Generic Global Model Development and Performance Evaluation

Part Six: Conclusions and Discussions

Chapter TWELVE
Conclusions and Discussions

Part Seven: References

Part Eight: Appendices

Tables

Figures

Part One

Background of Thesis

- Chapter One: ***Introduction***

  Chapter One provides an introduction of the thesis background. This includes: research motivation, research objectives, research questions, research scope, fundamental theory, originality of research and research design.

Chapter ONE

Introduction

1.1 Introduction

How can corporate performance be measured and forecasted? This question is of interest not only to company management but also to external stakeholders of the company, such as investors, potential lenders and policy makers. These players are continuously seeking the optimal solution to rationalize the decision-making process through performance evaluation.

There is no single approach to the evaluation of a company's performance, which is universally accepted. For example, a global survey of the top 500 largest firms was carried out by 'Fortune Magazine' in 2002. The main criterion of this survey is each company's sales. The ranking results for the global top five firms are illustrated in Table 1.1:

Table 1.1 Global Rankings of Top Five Largest Companies

Ranking of 2001	Company Name	Nationality	Sales in 2001 ($, Millions)	Profits in 2001 ($, Millions)	Ranking of Profit
1	Wal-Mart Stores	US	219,812	6,671	16
2	Exxon Mobil	US	191,581	13,520	1
3	General Motors	US	177,260	601	207
4	BP	Britain	174,218	8,010	6
5	Ford Motor	US	162,412	(5,453)	484

Source: Hjelt, P. (2002) Global 500 Companies, Fortune Magazine, 16 August, p. 54-60

Rankings show Ford Motor to be the fifth largest firm in the world. However, Ford Motor also had a big loss in 2001. Although Ford Motor had a good performance in terms of the scale of sales in 2001, it was not profitable over the same time period. This example shows that it is difficult to evaluate a company's performance only based on a single measure.

Corporate performance measurement is one of the key research topics in the *Management Accounting* domain. The primary objective of management accounting is to establish a decision support system for decision makers. A number of studies argued that incorporating both financial and non-financial performance indicators would provide a reliable framework for evaluating corporate performance (e.g. Amir and Lev, 1996; Behn and Riley, 1999). Numerous performance measurement systems were developed based on both financial and non-financial measures. One example is the *Balanced Scorecard* (Kaplan and Norton, 1996), which has been widely adopted for corporate performance assessment and strategic control in practice.

Despite providing frameworks for performance evaluation, few empirical studies examine the prediction power of these performance measurement systems, at least in the short term (Smith, 2005). As decision makers need to make decisions for future operations under uncertainty, they are more interested in predicting future performance rather than simply evaluating current performance. Models that forecast performance are therefore critical in the performance measurement domain.

Future corporate performance can be predicted through corporate credit risk evaluation. Credit risk is defined as the likelihood of borrowers (or creditors) not being able to meet their future obligations. There are several methods for predicting creditworthiness. For example, based on different types of data used, credit risk prediction models can be divided into two categories: *Market Based Model* and *Accounting Based Model*. Market based model, such as Moody's KMV model, is founded on Merton's (1974) option pricing framework. Accounting based model is developed using *Credit Scoring* techniques. In this research, the accounting based model is selected for final model construction, given the model's strong connection with performance measurement in the management accounting domain.

Credit scoring relies on a set of techniques derived from statistics or artificial intelligence domains, and in the present context, is applied to corporate credit risk assessment. The main objective of credit scoring is to establish a classification rule to distinguish the credit risks between 'healthy firms' and 'distressed firms'. Following a credit scoring process, each company is attributed a score based on certain specific measures. These credit scores enable the forecasting of a company's probability to face financial distress and assist performance prediction.

Currently, most financial distress prediction models can be criticized for their lack of theoretical groundwork for variable selection. Often, financial distress researchers select independent variables for model construction using previous studies. Such a variable selection method is limited for it fails to provide a holistic framework for research in financial distress prediction. In addition, measuring and predicting performance are interrelated issues that should be studied together. As a result, the following questions arise: Can we develop a performance measurement system based on a theoretical framework? Can we apply this theoretical framework to predict corporate default by using credit scoring techniques? These questions inspire me to explore further research.

1.2 Research Objective and Research Questions

With the aim of achieving the response to questions posed at end of Section 1.1, it is argued that a theoretical framework is required to address both performance measurement and performance prediction. Therefore, the main objective of this research is: ***Developing an Effective Corporate Performance Measurement and Prediction System***. This primary objective can be divided into three sub-objectives, as below:

1.2.1 Developing a Corporate Performance Measurement Framework

In order to construct a corporate performance measurement framework, a target industry and market should be defined for the purpose of determining an appropriate theory. Based on the fundamental theory, the next important task is to carry out a search for all potential performance measures, which could be financial, non-financial or external environmental variables. This search requires both primary and secondary data collection techniques.

1.2.2 Developing Financial Distress Prediction Models

The main tasks for developing a financial distress model include: sample selection criteria definition, relevant data collection, variables selection, modelling techniques determination, cross-validation technique selection and other issues related to modelling process.

1.2.3 Model Utility Evaluation

The research objective is not only to develop a system, but to construct an ***effective*** system. The main purpose of the model utility evaluation is therefore to prove this system is useful. A model's utility can be assessed through two criteria: model prediction power and practical applicability. An effective corporate performance measurement and prediction system should show sound performance in terms of both prediction power and practical applicability. Table 1.2 lists the relevant research questions based on the research objectives discussed above.

Table 1.2 Research Questions

1. Primary Research Question

    - How can corporate performance be measured and forecasted?

2. Sub Questions

    - What are the target industry and target market?
    - What would be the appropriate theory for model construction?
    - What would be the main considerations for evaluating a company's performance?
    - What would be the appropriate set of performance measures?
    - What credit scoring techniques will be employed in this research?
    - How can we evaluate the classification utility of the prediction model?
    - If the model's prediction utility is sound, can we apply this model to the real world?
    - How is the model's practical applicability?

1.3 Research Scope

A single industry is chosen to be the target of research. It has been argued by a number of authors that generic models for all industries tend to lack the ability to deal with specific industrial sectors. For example, Williams and Goodman's (1971) study showed that companies could be classified into different industries given their financial ratios. In addition, Bowen et al. (1982) pointed out that different industries have different financial structures (or leverage). Furthermore, Mensah (1984) also argued that even if in the same economic environment, different financial distress prediction models will have better performance when applied separately to different industries. These results imply that every industry has its distinctive financial characteristics and applying the same analysis to different industries may lead to irrelevant conclusions.

The retail industry is selected to illustrate the workings of the model. Partial reason for the choice is the researcher's interest in retail sector and his understanding of this field thanks to previous working experience in this industry. Moreover, the retail industry is a good candidate for research on corporate risk, due to the market structural changes in the retail environment, retail risk evaluation has become increasingly important. Dawson (2000) argued that one of the important retail research issues for the next five years is *Retail Risk Assessment and Evaluation.* There exist studies on risk assessment within a general business context (e.g. Ansell, 1992), but few on risk measurement within the retail industry. As a result, Dawson (2000) suggested that one potential retail research direction is to assess retail risk based on a theoretical framework.

The USA retail sector was chosen as the target market because of the clear definitions and reporting of financial distress through Chapters 7 and 11. Moreover, compared with other countries, the US market has the advantage of data completeness and sufficiency, especially for data from distressed firms.

1.4 Fundamental Theory

The research is based on the Hunt's (2000) *Resource-Advantage (R-A) Theory of Competition.* The R-A theory was selected as the primary theory due to its capability to describe the dynamic process of retail competition. Moreover, the fundamental premises of R-A theory are highly related to the real retail competition environment. Finally, R-A theory considers corporate competition advantage not only based on a company's internal resources, but also takes into account the influences from the external environment. It provides a more complete blueprint for research framework construction.

1.4.1 Foundational Premises

Compared with the traditional perfect competition theory, the foundational premises of R-A theory provide a more realistic description of the competition process (Hunt and Arnett, 2001). For example, in perfect competition theory, the primary resources in a company are those quantifiable and easily differentiated factors of production: land, labour and capital. Perfect competition theory tends to ignore a number of significant intangible factors such as entrepreneurship or a company's relationship with its suppliers. As such, it fails to reflect the real world. In contrast, R-A theory regards intangible factors, as important resources that can enhance a company's market position (Hunt and Arnett, 2003).

The premises of R-A theory are also closer to the retail competition environment. For example, R-A theory assumes that customers have imperfect and costly information. This premise indicates that why retailers value brand development, because they know customers wish to reduce product searching costs. Dawson (2000) pointed out that changing the nature of brands is one of the main challenges for future retail management. Nowadays, brand is not only a label of product, but also a technique to create various added values, such as, allowing higher margin on private label products, or protection from economic downturn (Fitch Ratings, 2000).

1.4.2 Dynamic Competition Process

Based on the foundational premises of R-A theory, Hunt and Morgan (1997) pointed out that if a company has a comparative advantage (or disadvantage) in resources, it would also possess a comparative advantage (or disadvantage) in the marketplace. A comparative advantage in the marketplace ultimately leads to superior financial performance. As management is always seeking superior financial performance through innovation and learning, the R-A competition process is dynamic, evolutionary and in disequilibrium. Consequently, the competition process of R-A theory is useful for demonstrating the highly dynamic retail competition environment (Hasty and Reardon, 1997).

1.4.3 Performance Measures

Hunt (2000) divided a company's resources into seven categories: financial resources, physical resources, legal resources, human resources, organizational resources, informational resources and relational resources. In addition, Hunt and Morgan (1997) pointed out that the competition process is not only affected by a company's internal resources, but also by external environmental factors, including societal resources on which firms draw, societal institutions that structure economic actions, actions of competitors and suppliers, behaviours of consumers, and public policy decisions. Drawing on above, R-A theory provides a more complete foundation for retail model framework construction. Further details on R-A theory and its assumptions will be presented in Chapter Three.

1.5 Research Design

A series of research strategies should be determined in order to achieve the research objectives. This section will focus on one main question: 'How does the researcher plan this research?'

1.5.1 Positivism versus Interpretivism

Weber (2004) compared the differences between *Positivism* and *Interpretivism* in terms of seven metatheoretical assumptions: ontology, epistemology, research object, research method, theory of truth, validity and reliability. He pointed out that positivists prefer to use scientific or quantitative research methods and they believe that research findings are based on the *'objectivity'* and independent of the human mind. Hence, research findings can be applied to other equivalent environments. However, interpretivists argue that reality and personal beliefs are not separable, for reality is too complex to describe using specific rules. In other words, interpretivists prefer to use *'subjective'* ways or qualitative approaches to do research so as to grasp deep meanings. The differences between positivism and interpretivism are summarized in Table 1.3:

Table 1.3 Differences between Positivism and Interpretivism

Assumptions	Positivism	Interpretivism
Ontology	Person and reality are separable	Person and reality are inseparable
Epistemology	Knowledge building in reality beyond the human mind	Knowledge building reflects people's living experience
Research Object	Research objects are independent of the researcher	Research objects are products of researcher's life-worlds
Research Method	Laboratory experiments, surveys, statistics...etc	Case studies, phenomenographic studies...etc
Theory of Truth	One-to-one mapping between researcher's statement and reality (beyond the human mind)	The truth depends on researcher's interpretation through the researcher's living experiences
Validity	Data can truly measure reality	Researcher's knowledge claims are defensible
Reliability	Data is reliable and can be replicated	Research is reliable, if researchers can demonstrate interpretive awareness

Source: Weber, R. (2004) The rhetoric of positivism versus Interpretivism: A personal view, MIS Quarterly, 28, 1, p.iii-xii

Review of previous literature on financial distress prediction indicates that most studies construct default prediction models based on positivism, especially in terms of variables selection. For example, most previous studies selected variables mainly from other previous studies or secondary data sources. In other words, these studies did not examine the validity of the selected variables based on the interpretivist viewpoint, since they believe these selected variables can truly evaluate reality. Obviously, this conclusion is naïve and does not reflect the real world.

On the other hand, if financial distress prediction studies only adopt the interpretivist viewpoint, the practical applicability of the model may be equally limited. The main reason would be that the model may not be replicable in different markets or industries, since the truth only depends on interpretation through the researcher's lived experiences.

In sum, both positivism and interpretivism have their own advantages and limitations and the choice of research philosophy depends on the research motivation and objective. In this research, the model framework is founded on both positivism and interpretivism. Performance measures were selected not only from secondary data sources, but also from primary data sources. Thus, interviews with interest parties were carried out with the intention of enhancing the findings from the review of secondary materials. In addition, a survey was also carried out in order to obtain more insights from stakeholders. This also allows a robust examination of the model framework.

1.5.2 Induction versus Deduction

Induction is the approach by which a theory is generated by comparing different phenomenon in reality and finding a common rule. *Deduction* is the method of applying theory to real world with the aim of examining the utility of the theory or exploring new insights from the original theory. Many social science studies start from developing a theory by using the inductive approach and then assess the utility of the theory through the deductive method. If the utility of theory is not sound, then modifications are necessary for the original theory by using the inductive approach again. The same process can be continued, if further modifications are necessary (Fielding and Gilbert, 2000).

This research is based on both inductive and deductive approaches. The model framework is constructed in an inductive way, by applying concepts from literature review as well as from interviews with experts. 170 potential retail performance measures are obtained for the model in this manner. After the variables collection process, a survey is carried out in order to ensure the validity and reliability of these performance measures. Subsequently, based on the fundamental theory, all qualified performance measures are divided into two variable groups: *Internal Resource Group* and *External Factor Group*. This provides a blueprint for final model framework construction.

After developing the framework, the next question will be: 'Is the framework useful?' This question requires more exploration and reflects the deductive approach. Default prediction models are developed to examine the utility of the framework. A straightforward way to evaluate a default prediction model's performance is its predictive power. In other words, the more accurate the prediction, the better the model's performance and hence, the better the framework's utility. Drawing on this insight, the design of this research includes three main elements: ***Theory***, ***Framework*** and ***Model***. The relationship among these three elements can be expressed as three concentric circles in Figure 1.1:

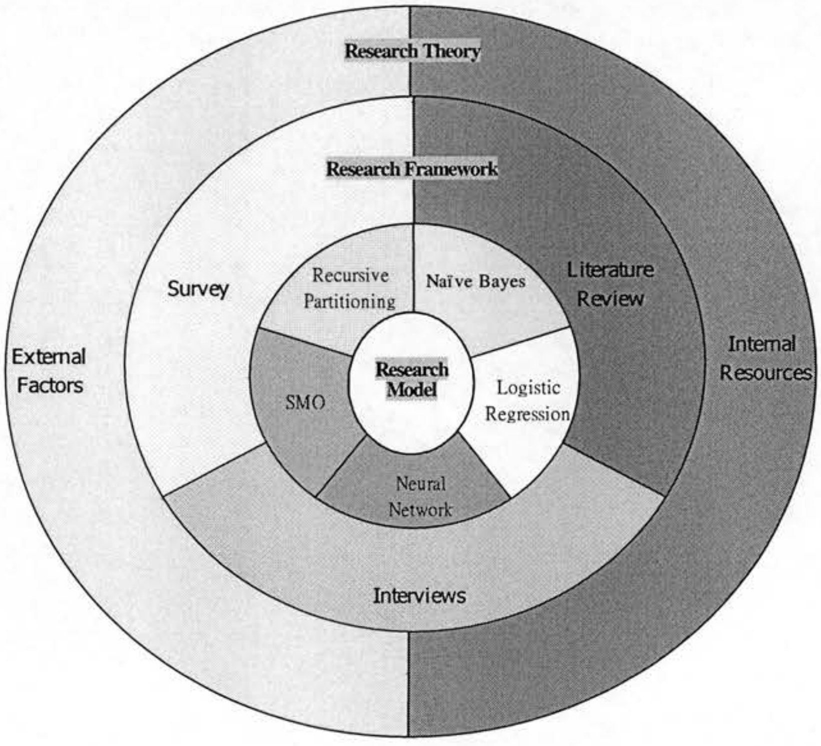


Figure 1.1 Research Design

The outer concentric circle is the research fundamental theory: *Resource-Advantage Theory of Competition*, which is based on both internal resources and external factors as discussed in Section 1.4.3. The second concentric circle indicates the model's framework, which is constructed by inductive research design through literature review, interviews and survey. From this, a large number of potential variables are found. Attempts to consider all conceivable performance measures bring too many variables in a model, causing overfitting and, hence, poor prediction. The number of variables needs to be reduced by using *'Cluster Analysis'* to eliminate the influences from outliers, *'Univariate Analysis'* to select candidate variables, *'Principal Component Analysis (PCA)'* or *'Stepwise Regression Approach'* to identify primary principal components or key variables.

The smallest concentric circle represents corporate default prediction model. The development of the model is based on the deductive approach using desk-based research approach. Five credit scoring techniques are selected for modelling purposes: *Naïve Bayes*, *Logistic Regression*, *Recursive Partitioning*, *Artificial Neural Network*, and *Sequential Minimal Optimization* on a sample of 195 healthy and 51 distressed firms over five time periods: 1994-1998, 1995-1999, 1996-2000, 1997-2001 and 1998-2002. This enables a comparative analysis from a time-scale viewpoint.

The corporate default prediction model is evaluated in terms of both model prediction power and practical applicability. *Classification Accuracy Rate* and *the Area under the Receiver Operating Characteristics Curve (AUROC)* are employed to evaluate the model prediction power. Practical applicability of the model is assessed by two approaches. First, rankings from this research are compared with those from a standard rating system—in this case *Moody's Credit Rating*. It is assumed that the higher the degree of similarity between the two sets of rankings, the better the developed model's practical applicability.

Another method to assess the model's practical applicability is to apply the original model to different markets and European and Japanese markets are selected for model application. This allows an international comparison based on the model's prediction performance. In spite of this, the development of a general global system for a single industry is critical, as model construction for individual country systems would be too time-consuming and costly. In this research, a composite model is constructed by combining data from US, European and Japanese markets.

From Figure 1.1 above, it can be concluded that if default prediction models have good performance in terms of both prediction utility and practical applicability, then the usefulness of the research framework and the fundamental theory can be confirmed. Moreover, it also implies that the theoretical groundwork for variable selection is critical in the financial distress prediction domain. Finally, the primary research objective: ***'Developing an Effective Corporate Performance Measurement and Prediction System'*** can be achieved.

1.6 Originality of Research

In the introduction section, the discussion centred on the existing gap between performance measurement and default prediction models in previous research. On the one hand, performance measurement systems are based on a theoretical framework but few empirical studies examine their prediction ability. On the other hand, default prediction models have the ability to provide a platform for forecasting company performance but most lack theoretical framework and fail to incorporate industrial viewpoints for variable selection. Therefore, an original endeavour in this thesis is to fill the gap between previous performance measurement systems and performance prediction models as well as to overcome the relevant limitations.

Unlike most previous financial distress prediction studies, this research developed a theoretical framework for model construction. Moreover, this research has taken into account subjective assessment through obtaining insights from context in order to ensure the practical applicability of the final default perdition model. The goal of filling the gap in previous studies is reached by constructing a model framework using Hunt's (2000) Resource-Advantage Theory of Competition and data from interviews and performance surveys with retail companies.

This research incorporated a much larger number of variables (original number: 170 variables) than in most previous studies. Considerations include qualitative variables, external environmental variables and cash flow structure variables. Moreover, most previous studies select external environmental variables only based on macroeconomical variables. However, this research also selects variables related to political, social-cultural and technological environments.

The work also explores the performance of several methodologies for predicting default: Naïve Bayes, Logistic Regression, Recursive Partitioning, Artificial Neural Network, and Sequential Minimal Optimization (a form of Support Vector Machine). In particular, the application of *Sequential Minimal Optimization (SMO)* is an innovation in the area of financial distress and business default predictions. It allows a performance comparative analysis among different credit scoring techniques to be carried out. Given the importance of developing techniques that can be easily interpreted, graphical representation of predictions are also explored.

1.7 Thesis Overview

Based on the research objectives, the main body of the thesis can be divided into six parts: *Background of Thesis*, *Review of the Literature*, *Research Framework Development*, *Default Prediction Model Construction*, *Model Utility Evaluation* and *Conclusion*. The thesis structure can be illustrated as follows:

1.7.1 Part One: Background of Thesis

Chapter One provides an introduction of the thesis background. This includes: research motivation, research objectives, research questions, research scope, fundamental theory, originality of research and research design.

1.7.2 Part Two: Review of the Literature

The review of previous studies on the development of performance measurement systems and corporate financial distress prediction models will be presented in

Chapter Two. The goals are to illustrate the main reason behind the research motivation and questions as well as to introduce the key issues related to performance measurement and default prediction research.

1.7.3 Part Three: Research Framework Development

Part Three includes three chapters (from Chapter Three to Chapter Five). *Chapter Three* looks at the fundamental R-A theory, as well as the basic premises of the theory structure and the implementation process. The main purpose is to provide a blueprint for research framework development.

Based on R-A theory, *Chapter Four* then focuses on the construction of the model framework. This involves literature review and interviews with practitioners in connection with the selection of retail performance measures. A pilot study of interview will also be presented in this chapter.

In *Chapter Five,* a survey analysis is presented. The survey provides real-world insights regarding the importance of performance measures. The primary purpose is to conduct a robust examination of the model framework. A series of comparative analysis are also carried out based on different countries, different retail management functions and different retail formats. Finally, a case study will illustrate the application of R-A theory.

1.7.4 Part Four: Default Prediction Model Construction

Part Four consists of Chapters Six and Seven. *Chapter Six* reviews the sample selection criteria, the data arrangement as well as the selection procedures for key variables, such as principal component analysis and stepwise regression procedures. The selected key variables and principal components will also be described.

Chapter Seven then concentrates on the modelling procedure, including an introduction of the selected credit scoring techniques and cross-validation approaches.

1.7.5 Part Five: Model Utility Evaluation

From Chapter Eight to Chapter Eleven, the main focus is on the performance assessment of the default prediction model. *Chapter Eight* will evaluate the default prediction models' performance by using predictive accuracy rates and AUROC. An introduction of the approaches for model performance measurement will also be presented. Furthermore, some key issues, such as detection of external influences, evaluation of types of error and the exploration of time series effects will also be addressed.

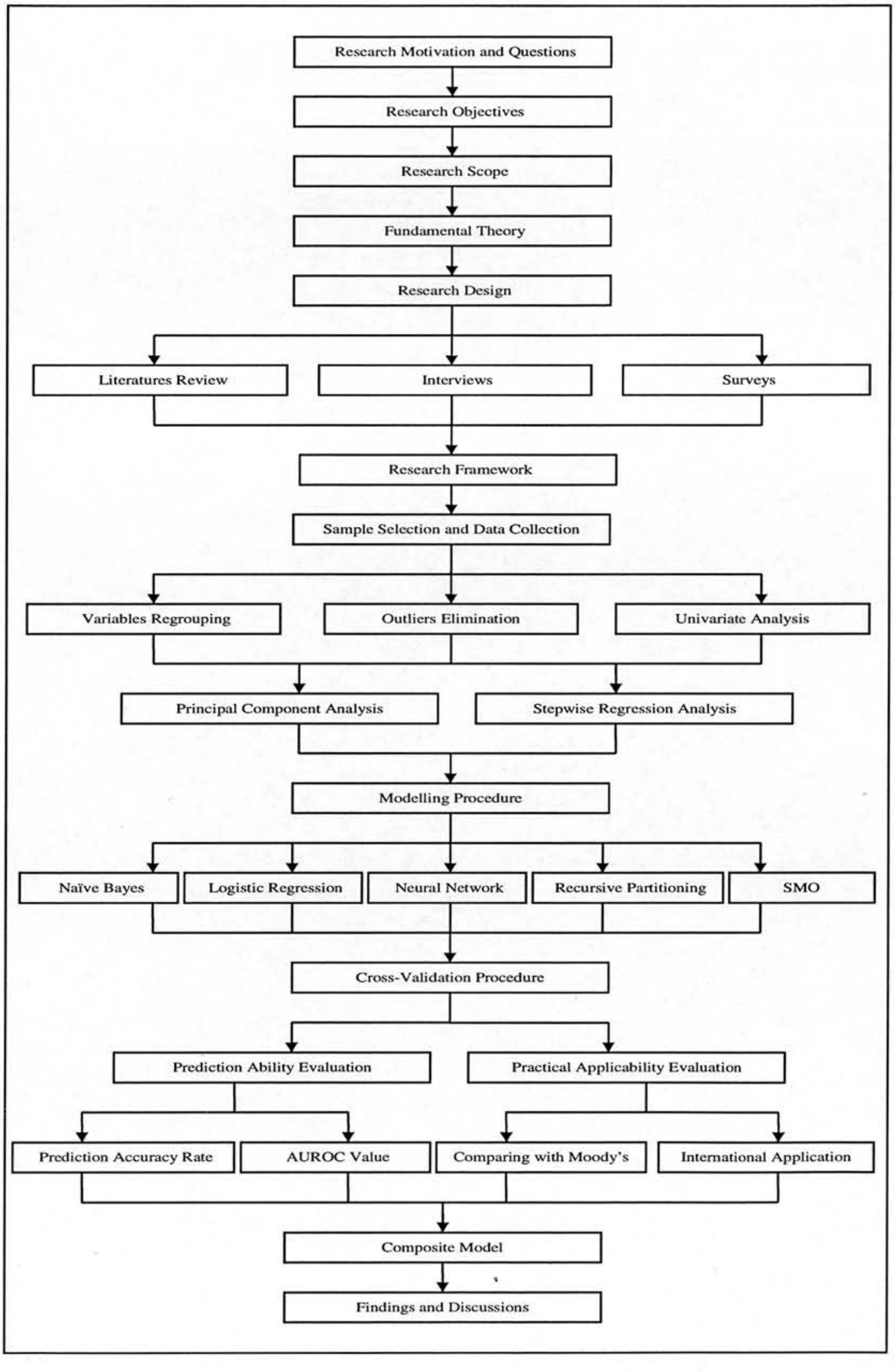
Chapters Nine and Ten follow with assessments of the practical applicability of default prediction models. In *Chapter Nine*, the assessment is carried out by comparing the PhD research models with Moody's credit ratings. Data collection for Moody's ranking and the techniques for comparison are also introduced. In *Chapter Ten,* the assessment of practical applicability consists of applying the model to other market datasets in different time periods. A new US dataset, a European dataset and a Japanese dataset are used. An international comparison analysis of the models' prediction performance is then carried out.

Chapter Eleven will focus on the development of a composite model based on combining data from US, European and Japanese markets. The prediction ability and practical applicability of the composite model will also be evaluated. Moreover, a comparative analysis between the composite model and the original USA model will also be conducted.

1.7.6 Part Six: Conclusions and Discussions

Chapter Twelve summarizes the findings, outlines the limitations and suggests possible future directions for research in the corporate performance measurement and default prediction domain. In addition, it illustrates the contributions to interested parties in this research area. The thesis structure is presented in Figure 1.2:

Figure 1.2 Research Structure



Part Two

Review of the Literature

- Chapter Two: ***Development of Performance Measurement and Prediction System***

The review of previous studies on the development of performance measurement systems and corporate financial distress prediction models will be presented in *Chapter Two*. The goals are to illustrate the main reason behind the research motivation and questions as well as to introduce the key issues related to performance measurement and default prediction research.

Chapter TWO

Development of Performance Measurement and Prediction System

2.1 Introduction

This chapter presents an initial review of previous studies on the development of performance measurement systems, followed by an introduction to the development of corporate default prediction models. A discussion of some issues related to corporate default prediction research will also be presented. The final section not only summarizes the key discussions but also explains the research motivation and the generation of research questions.

2.2 Development of Performance Measurement Systems

Corporate performance evaluation has become more difficult than in the last decades of the twentieth century. Part of the reason is the shift from industrial age competition to information age competition. Chandler (1990) pointed out that the main objective for companies in the industrial age (from 1850 to 1975) was to capture profits from the economies of scale and scope. Hence, corporate performance measurement was straightforwardly based on manufacturing efficiency.

However, in the information age, market competition is more severe and characterized by drastic changes in the operating environments. While firms in the industrial age focused on mass production with low-cost standardized products that are pushed to customers through the supply chain, the firms in the information age need to learn how to provide customized products to their target customers in order to satisfy their high-variety and low-volume demand without paying higher cost (Cooper and Kaplan, 1988). In fact, with the advances in manufacturing technologies, many researchers argued that the objective of 'mass-customization' is achievable; it is possible to give consideration to both production efficiency and demand variety (Pine II et al., 1993; Kotha, 1995). The development of the information age implies that even more considerations are needed for evaluating corporate performance.

Corporate performance measurement is one of the key research topics in the *Management Accounting* domain, given that the primary objective of management accounting is to establish a decision support system for decision makers. Smith (1997) pointed out that the traditional accounting performance measurement system focuses on 3 *Es*: ***Efficiency***—utilisation of equipment and workforce, ***Economy***—optimum use of resources and ***Effectiveness***—achievement of target outcomes, and this traditional system is developed in terms of the internal, quantitative, financial and accounting measures. Bromwich and Bhimani (1989) argued that the traditional management accounting only focuses on the 'factory floor' and hence, it cannot meet new market challenges in the information age. Obviously, this traditional system over-emphasizes productivity and fails to consider several important factors, such as non-financial measures (Johnson and Kaplan, 1987).

Amir and Lev (1996) examined the relationship between the valuation of cellular companies and a number of performance indicators, which included both financial and non-financial measures and they argued that on a stand-alone basis, financial measures were largely irrelevant for the security valuation, while some non-financial measures are highly relevant. Moreover, financial measures became relevant to the stock valuation when combined with the non-financial measures. It implies that non-financial indicators had better utility to measuring corporate financial performance. Moreover, Behn and Riley (1999) explored whether some non-financial indicators have the ability to predict financial performance of the U.S. domestic airline industry. Results based on one or two months of non-financial data showed that non-financial measures were useful to predict the financial performance of the target industry. Other studies have also indicated the utility of non-financial indicators to predict corporate financial performance – for example, Hughes II (2000), Banker et al. (2000) and Said et al. (2003).

The discussions above show that incorporating both financial and non-financial performance indicators would provide a reliable framework for evaluating corporate performance. In fact, a new term in management accounting, called *Strategic Management Accounting (SMA)* was developed in order to address the importance of qualitative aspects. SMA was first introduced by Simmonds (1981), and in his study, SMA was defined as:

> *"the provision and analysis of management accounting data about a business and its competitors for use in developing and monitoring the business strategy, particularly relative levels and trends in real costs and prices, volume, market share, cash flow and the proportion demanded of a firm's total resources."*

(Simmonds, 1981:26)

Roslender and Hart (2003) further defined SMA as:

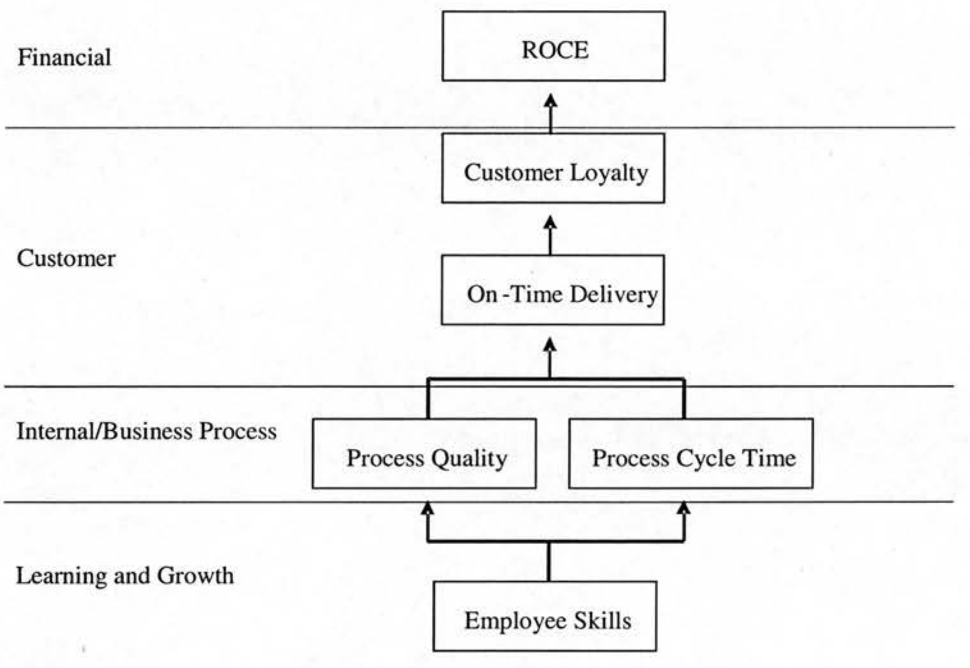
> *"SMA is a generic approach to accounting for strategic position. It is defined by an attempt to integrate insights from management accounting and marketing management within a strategic management framework"*

(Roslender and Hart, 2003:255)

From the definitions above, SMA's emphasis is placed on providing a strategic vision for traditional management accounting in order to facilitate management for decision-making (Bromwich 1988; Bromwich and Bhimani, 1989; Smith, 2005). So far, the concept of SMA has been extended to a number of corporate performance measurement systems, such as the *Strategic Cost Management* framework based on the value chain analysis (Shank and Govindarajan, 1992) and the *Balanced Scorecard* (Kaplan and Norton, 2001). Among these performance measurement techniques, the *Balanced Scorecard* has been widely adopted for performance assessment and strategic control in numerous companies.

The balanced scorecard focuses on four performance perspectives (*Financial Perspective*, *Customer Perspective*, *Internal-Business-Process Perspective* and *Learning and Growth Perspective*) and each perspective has its own objectives and corresponding performance measures (Kaplan and Norton, 1992). For example, one of the main objectives in the customer perspective is customer responsiveness and a corresponding performance measure would be on-time delivery (Smith, 2005). How does balanced scorecard work? Kaplan and Norton (1996) adopted the cause-and-effect relationships among these four perspectives to illustrate the implementation process of balanced scorecard. The cause-and-effect relationships can be demonstrated in Figure 2.1:

Figure 2.1 Cause-and- Effect Relationships



Source: Kaplan, R.S., Norton, D.P. (1996) The Balance Scorecard: Translating Strategy into Action, p31

Kaplan and Norton (1996) assumed that an increase in the value of a company's *Return on Capital Employed (ROCE)* indicates an improvement of profitability. One of the possible reasons to cause an increase of ROCE is sales expansion from existing customers. Hence, the customer loyalty is the primary driver for probability in this case. Customer loyalty can be established by many methods, such as on-time delivery, corporate brand or image. In addition, on-time delivery depends on the efficiency of the internal process, such as high process quality and short process cycle time. In turn, the effectiveness of internal process is based on the high quality human resource through training and learning.

Although the balanced scorecard has shown its practical applicability in the real world, the cause-and-effect relationships are still debatable in academia. Reinartz and Kumar (2002) examined the relationship between customer loyalty and accounting profitability by using customer databases at four companies and argued that no strong direct correlation between customer loyalty and customer profitability exists. Furthermore, Ittner et al. (2003) pointed out that even if customers have high satisfaction with the balanced scorecard, there is no relationship with the improvement of economic performance. A number of studies also argued that there is no obvious association between customer satisfaction and financial performance (Ittner and Larcker, 1998; Norreklit, 2000; Garland, 2002).

Moreover, as these performance measurement systems focus on risk assessment within a general business context, they tend to lack the ability to deal with specific industrial sectors. A number of studies argued that different industries have different financial characteristics (Williams and Goodman, 1971; Gupta and Huefner, 1972; Bowen et al., 1982; Mensah, 1984) and applying the same analysis to different industries may lead to irrelevant conclusions.

For instance, suppose investors are interested in the goods' unsalable risk of two firms from the construction and retail industries, and decide to use inventory turnover to measure such risk. The construction company might need six months to sell a product, but the retail company might sell a product in ten minutes. The inventory turnover of the construction company would be much lower than the retail company's. Does this mean that the goods' unsalable risk is lower in the retail company? Clearly, there is no absolute answer in such a situation, since the companies observed are from different industries.

Finally, despite providing frameworks for performance evaluation, few empirical studies examine the prediction power of these performance measurement systems, even in the short term (Smith, 2005). As decision makers need to make decisions for future operations under an uncertainty environment, they are more interested in future performance prediction than current performance evaluation. Drawing on this insight, performance forecasting models are also crucial in the performance measurement domain.

2.3 Development of Default Prediction Model

Credit risk evaluation is one of the approaches that can be used to predict a company's performance. Credit risk is defined as the likelihood of borrowers (or creditors) not meeting their future obligations. These borrowers can be individuals, companies or sovereign governments. *Healthy* borrowers have lower credit risk than *distressed* borrowers, and hence, show better performance.

Corporate credit risk can be evaluated by a variety of modelling techniques. Caouette et al. (1998) classified credit risk modelling techniques into five categories: Econometric Techniques, Computer-based Systems, Optimization Models, Rule-based Systems (or Expert Systems) and Hybrid Systems, which are summarized in Table 2.1:

Table 2.1 Credit Risk Modelling Techniques

Category	Techniques
Econometric Techniques	Discriminant Analysis, Multiple Regression, Logit analysis, Probit Analysis, Survival Analysis
Computer-based Systems	Neural Network
Optimization Models	Mathematical Programming
Rule-based or Expert Systems	Try to mimic in a structured way the process that an experienced analyst uses to arrive at the credit decision.
Hybrid Systems	Using direct computation, estimation and simulation are driven in part by a direct causal relationship, the parameters of which are determined through estimation techniques, such as, Moody's KMV model.

Source: Modified from Caouette et al., (1998) Managing credit risk: The next great financial challenge, p.104-105

Based on different types of data used, corporate default prediction models can also be divided into the *Market Based Model* and *Accounting Based Model*. In this research, the accounting based model (or credit scoring model) was selected for final model construction, since it has strong connection with performance measurement in the management accounting domain. The next section outlines the market based model, followed by a detail illustration on the development of the accounting based model.

2.3.1 Market Based Model

Market based model, such as Moody's KMV model (Crosbie and Bohn, 2005), is founded on the Merton's (1974) option pricing framework. The basic assumption of this approach is that if a company's market value of assets is equal or less than its book value of liabilities, then the company can be considered as to have defaulted. The underlying logic is that the company will not have the ability to face its future obligations even if it sells all its assets in the market. The distance between the market value of assets and book value of liabilities—*Distance-to-Default (DD)*— indicates the likelihood of a company to face financial distress and DD can be expressed in the following function:

$$DD = \frac{V_A - L}{V_A \times \sigma_A} \tag{2.1}$$

where V_A is the market value of assets, L is the book value of liabilities, σ_A is the asset volatility.

In order to calculate the distance-to-default, the first step is to estimate the V_A and σ_A by using Merton's (1974) framework for a company's equity can be regarded as a call option on the underlying assets of the firm. Merton's (1974) framework is illustrated as follows:

$$V_E = V_A N(d_1) - De^{-\tau r} N(d_2) \tag{2.2}$$

$$\sigma_E = \frac{V_A}{V_E} \Delta \sigma_A \tag{2.3}$$

where

$$d_1 = \frac{\ln(\frac{V_A}{L}) + (r + \frac{\sigma_A^2}{2})\tau}{\sigma_A \sqrt{\tau}} \tag{2.4}$$

$$d_2 = d_1 - \sigma_A \sqrt{\tau} \tag{2.5}$$

V_E is the market value of equity (or market capitalization), r is the risk free interest rate, τ is the maturity of the liabilities.

Given that r and L are public information and V_E and σ_E can be easily calculated by using the stock price data in the equity market, V_A and σ_A can be estimated from Function (2.2) and Function (2.3). In addition, the value of distance-to-default can be further calculated by using the Function (2.1) to express the likelihood of default.

Although the market-based model is useful to predict a company's credit risk, its assumptions do not always reflect reality. For example, the market-based model assumes that default occurs as soon as a company exhausted its assets. However, default usually occurs *before* a company exhausted its assets (Longstaff and Schwartz, 1995). In addition, the market-based model assumes *'Perfect Market'*, which means that no transaction costs or taxes exist and that lending and borrowing interest rates are identical in the exchange market.

Furthermore, the risk-free interest rate remains constant over time and term structure effects will not have any impact on the pricing process. Finally, the model assumes that all liabilities will be due simultaneously, regardless of the type of debt. In reality, a company's debt structure is more complex and usually includes various types of liability, such as corporate bonds or business notes. The assumptions are difficult to implement in practice.

A number of studies attempted to enhance the Merton's framework in order to address these limitations. Regarding the assumption of the fixed interest rate, many studies took into account floating interest rates (or stochastic interest rates) in the pricing model (Shimko et al., 1993; Longstaff and Schwartz, 1995). Furthermore, a range of articles considered various debt structures in their research in order to address the complexity of a company's financial structure (Jones et al., 1984; Collin-Dufresne et al., 2001).

2.3.2 Accounting Based Model

2.3.2.1 Overview

The accounting-based model, such as the Beaver's (1966) univariate model, Altman's (1968) Z-score model and Ohlson's (1980) logistic regression model, is developed using *Credit Scoring* techniques with financial measures, non-financial measures, macro-economical measures, or other relevant measures. According to the definition in Thomas et al. (2002), credit scoring can be defined as:

> *"the set of decision models and their underlying techniques that aid lenders in the granting of consumer credit. These techniques decide who will get credit, how much credit they should get, and what operational strategies will enhance the profitability of the borrowers to the lenders."*

(Thomas et al, 2002:1)

From the definition, credit scoring can be regarded as a collection of techniques, which are derived from statistics or artificial intelligence domains, and used for individual, corporate or government credit risk assessment. In connection with the corporate credit risk evaluation, the main objective of credit scoring is to establish a classification rule to distinguish the credit risk between 'healthy firms' and 'distressed firms' so as to assist lenders to make loan decisions. Following a credit scoring process, each company is attributed a credit score based on some specific measures. These credit scores enable the forecasting of a company's probability to face financial distress and hence, achieve the goal of predicting performance. This section will introduce different accounting based modelling techniques in terms of different time periods. Moreover, the advantages and shortcomings of each modelling technique will also be discussed.

2.3.2.2 Univariate Analysis

Beaver (1966) was a pioneer in financial distress prediction research with a number of authors following his work. In his study, he choose distressed firms from the *Moody's Industrial Manual* and selected healthy firms based on a paired-sample design in order to eliminate the influences from the variety of size and industry. 30 financial measures were selected based on the following criteria: 1) popularity: frequent appearance in the previous literatures, 2) utility: successful performance in one of previous studies, and 3) cash flow framework.

For cash flow framework, Beaver (1966) viewed a company as a reservoir supplied by the liquid-asset inflows and drained by the liquid-asset outflows. The default can be identified based on the likelihood that the reservoir will be exhausted. Beaver argued that a company will have lower probability to face financial distress, if: 1) the reservoir is larger, 2) the net liquid-asset flow from operations is larger, 3) the amount of debt held is smaller, and 4) the fund expenditures for operations are smaller. Based on the cash flow framework, Beaver then carried out three different univariate analyses—profile analysis (comparison of mean values), dichotomous classification test and likelihood ratio analysis—in order to examine the predictive characterises and utility of each variable.

From profile analysis, Beaver (1966) compared the mean values among 30 financial ratios and found that distressed firms had several characteristics, such as, lower cash flow, smaller reservoir of liquid assets and more debt hold, which are consistent with the basic assumptions of the framework. With regards to the dichotomous classification test, Beaver first determined an optimal cut-off point distinguishing healthy from distressed companies for each financial ratio. He then applied these cut-off points to a holdout sample to evaluate each financial ratio's prediction power by using the univariate dichotomous classification test. He argued that the variable with the best predictive performance is the *Cash Flow to Total Debt Ratio (CFTD)*. The prediction accuracy of CFTD rate is 87% one year before financial distress and is 78% five years before financial distress.

Finally, Beaver (1966) conducted an analysis of likelihood ratios based on the *Bayesian* approach. He argued that the default prediction problem could be regarded as a problem of evaluating the probability of financial distress conditional upon the value of a specific financial ratio. He further pointed out that financial ratios can provide useful information for predicting default, since the likelihood ratios still present high values even five years prior to financial distress.

Univariate analysis is limited in the evaluation of a firm's performance, since it is difficult to use only one single measure to describe the performance in a multidimensional firm. However, prior to construct a multivariate model, it is still useful to carry out a univariate analysis for the purpose of variable selection, as not every variable has good discriminating utility (Hosmer and Lemeshow, 2000).

2.3.2.3 Multiple Discriminant Analysis

Altman (1968) suggested using Multiple Discriminant Analysis (MDA) to develop a Z-score bankruptcy prediction model. He selected distressed firms based on Chapter X of the National Bankruptcy Act in the manufacturing industry. Like Beaver (1966), Altman adopted the pair-sample design to choose healthy firms. Overall, 33 healthy and 33 distressed firms were selected from 1946 to 1965. Regarding variable selection, Altman selected five key variables from 22 potential variables in terms of four criteria: 1) statistical significance, 2) evaluation of inter-correlations, 3) predictive accuracy and 4) judgement of the analyst.

MDA is a multivariate statistical technique using a linear combination of independent variables. Through maximizing the discriminant criterion, which is the variance between the healthy group and distressed group divided by the variance within each group, weights will be attributed to the independent variables and the Z-score function can be generated, (Fisher, 1936; Thomas et al. 2002; Chou, 2002). Altman's (1968) Z-score function is expressed in Function 2.6:

$$Z = 0.012\, X_1 + 0.014\, X_2 + 0.033\, X_3 + 0.006\, X_4 + 0.999\, X_5 \qquad (2.6)$$

where X_1 = Working Capital / Total Assets, X_2 = Retained Earnings / Total Assets, X_3 = Earnings before Interest and Taxes / Total Assets, X_4 = Market Value Equity / Book Value of Total Debt, X_5 = Sales / Total Assets.

In Altman's (1968) research, a Z-score cut-off point (2.675) was determined in order to classify healthy and distressed firms. The results showed that the Z-score model had sound prediction performance one year and two years before financial distress, but did not indicate good prediction utility three to five years before financial distress. In order to expand the application scope of the Z-score model, Altman (1995) modified the original Z-score model to apply to the private firms by changing the variable *Market Value Equity/Book Value of Total Debt* to *Book Value Equity/Book Value of Total Debt.*

Altman et al. (1995) further revised the Z-score model with the intention of applying the Z-score model to a non-manufacturing industry. Finally, Altman et al. (1977) developed a ZETA model by using seven variables with the purpose of enhancing the original Z-score model. They argued that the new ZETA model has sound predictive performance with the accuracy rate over 90% one year before financial distress and over 70% five years prior to default.

In the UK market, Taffler (1983) also employed MDA with 80 potential variables to develop an UK-based Z-model and results indicated that the UK-based Z-model presented a high prediction accuracy performance. After Altman's (1968) research, a number of studies also used MDA to predict firm's default in different markets or different industries, including Deakin (1972), Blum (1974), Libby (1975), Altman et al. (1977), Taffler (1982, 1984), Pantalone and Platt (1987), Betts and Belhoul (1987) and Piesse and Wood (1992).

Although previous MDA models presented good prediction performance, MDA still has a number of potential statistical problems. Eisenbeis (1977, 2004) grouped these problems into eight categories, which are illustrated in Table 2.2:

Table 2.2 Statistical Problems of the Application of Discriminant Analysis

1. The distributions of the variables
2. The group dispersions
3. The interpretation of the significance of individual variables
4. The reduction of dimensionality
5. The definitions of the groups
6. The choice of the appropriate *a prior* probabilities and /or costs of misclassification
7. The estimation of classification error rates
8. The selection of the analysis samples.

Source: Eisenbeis (1977) Pitfalls in the application of discriminant analysis in business, finance and economics, Journal of Finance, 32 (3) p.875-900; also see Thomas et al. (2004) Readings in Credit Scoring Foundations, Developments, and Aims, Oxford University Press, p.24

For example, MDA assumes that the covariance matrices of two populations are identical and both populations need to be described by multivariate normal distribution. Clearly, these assumptions do not always reflect the real world. Deakin (1976) argued that even after performing the normality transforming process, financial ratio data do not follow normal distribution. Moreover, Hamer (1983) evaluated the sensitivity of financial distress prediction models in terms of four different variable sets from previous research (Altman, 1968; Deakin, 1972; Blum, 1974; Ohlson, 1980). She pointed out that the covariance matrices in each variable set were statistically different. Although the problem of unequal covariance matrices can be solved by employing the quadratic discriminant function, some studies still indicate that the linear discriminant analysis shows better performance than the quadratic discriminant analysis (Altman et al., 1977; Hamer, 1983).

2.3.2.4 Conditional Probability Model

To avoid the limitations of MDA, Ohlson (1980) was the first to apply the conditional probability model and in particular, the *Logistic Regression* model, to bankruptcy prediction research. Unlike MDA, the logistic regression model requires neither multivariate normality nor the equality of covariance matrices of two populations. By logit transformation on odd ratio function, the logistic model can be linearized and used to solve classification problems. A logistic function can be expressed as follow:

$$g(x) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n = B \times X^T \qquad (2.7)$$

where $\pi(x)$ is the logistic function,

$$\pi(x) = \frac{1}{1+e^{-(B \times X^T)}} = \frac{e^{B \times X^T}}{1+e^{B \times X^T}} \qquad (2.8)$$

Like Altman's (1968) study, Ohlson (1980) selected distressed firms based on a legally viewpoint: either Chapter X or Chapter XI. In addition, the selected distressed firms had to be publicly listed industrial[1] companies from 1970 to 1976. Overall, the sample size for distressed firm was 105 and 2028 for healthy firms. Nine variables were selected based on the popularity in previous studies and results showed that four basic performance measures were statistically significant: 1) *Size measure*: log (total assets/GNP price-level index), 2) *Leverage measure*: total debts/total assets, 3) *Performance measures*, such as, net income/total assets and 4) *Liquidity measures*: such as, working capital/total assets.

Ohlson (1980) also pointed out that the timing of data collection for distressed firms is crucial. Since the auditing process is time-consuming, the date of financial statements releasing is usually after the date of the financial year-end. It is possible

1 The utilities, transportation firms and financial service companies were excluded.

that a company files for bankruptcy between the date of financial statements releasing and the date of the financial year-end. Drawing on this insight, the most recently financial statement before financial distress may only be available after bankruptcy is filed. The major problem is that the financial reports *after* financial distress would usually include the adjustments from auditors in light of the bankruptcy filing and they are not comparable with the normal year end financial reports. To solve this problem, Ohlson (1980) suggested that financial statements prior to the financial distress year should be viewed as the last report if firms file for bankruptcy after the date of financial year-end, but prior to the date of financial statements releasing. Following Ohlson's (1980) study, Mensah (1983), Casey and Bartczak (1985), Gentry et al. (1985), Keasey and McGuinness (1990), Tennyson et al. (1990), as well as Ward (1994), also employed the conditional probability models to predict financial distress.

Although logistic regression does not suffer from the limitations of MDA, Tabachnick and Fidell (2000) pointed out that if the assumptions regarding the identical covariance matrices and multivariate normal distribution are met, MDA is likely to be more efficient than logistic regression. Moreover, like all the regression functions, the problem of multicollinearity still exists in logistic regression.

2.3.2.5 Recursive Partitioning

In the mid-1980s, *"Recursive Partitioning Analysis (RPA)"* or *"Decision Tree"* was introduced in the financial distress prediction research area, (Marais, et al., 1984; Frydman et al., 1985; Carson and Hoyt, 1995). RPA is a non-parametric technique and does not suffer the limitations from MDA or conditional probability model. Although Fisher's (1936) linear discriminant method is often viewed as the oldest classification technique, Hand (1997) argued that the basic idea of RPA is very straightforward, and hence the oldest conceptually.

RPA can be regarded as a stepwise procedure. The first step is to select an independent variable as the best discriminator and to decide on a cut-off point value based on the lowest expected misclassification cost. Based on the cut-off point, the second step is to divide both healthy and distressed firms into two sub-nodes. The third step is to select another (or the same) discriminator and further partition the healthy and distressed firms into another two sub-nodes. The same process can be continued, if further splitting is necessary. Thomas et al. (2002) mentioned two reasons to stop the partitioning process. First, if the number of samples in a node is too small, then further partitioning is not appropriate. Second, if the classification results between the old and new nodes do not have significant differences, then it is not necessary to split the old node.

Frydman et al. (1985) developed a RPA default prediction model by using a sample of 58 distressed firms and 142 healthy firms from 1971 to 1981. 20 financial variables were selected for model construction based on three previous studies: Altman (1968), Deakin (1972) and Altman et al. (1977). The main objective is to compare the predictive performance between RPA and MDA and results showed that RPA has better classification ability then MDA. However, RPA does not always show superior performance. Marais et al. (1984) examined several key issues of experimental design regarding the classification of bank loans by using conditional probability model (polytomous probit model) and RPA. They pointed out that the polytomous probit model and RPA have essentially equivalent classification ability.

Although RPA is not affected by the assumptions required for MDA or conditional probability model, it still has some limitations. For example, it is difficult to explore the relative importance of each variable by using RPA. Frydman et al. (1985) argued that RPA is similar to the forward stepwise approach for variable selection. As long as a variable is selected as the discriminator, this variable will be considered for the next selection. Moreover, it is highly possible to select the same variable to be the discriminator again. Thus, the contributions of each variable to the dependent variable are ambiguous in RPA. Another major problem relative to RPA is overfitting: the continuous partitioning process is likely to encourage one misclassified case in the terminal node (Thomas et al., 2002). The overfitting problem can be overcome by a *'Cross-Validation'* procedure. This concept will be illustrated in Chapter Seven.

2.3.2.6 Expert System

Expert system or the Rule-based system was developed to predict corporate bankruptcy in the late 1980s. Thomas et al. (2002) defined an expert system as being based on a set of rules that will imitate an expert's behaviour in decision-making. In addition, Caouette et al. (1998) defined an expert system as a computer-based decision support system in the artificial intelligence domain. Overall, an expert system can be regarded as a knowledge-based computing system that can provide recommended suggestions to facilitate users to make decisions.

Elmer and Borowski (1988) developed an expert system to explore the bankruptcy issues on saving and loan (S&L). This system is based on a collection of rules gathered from government, industry analysts and other sources relative to S&L financial analysis. The framework of the expert system is founded on the *Capital, Assets, Earnings and Liquidity (CAEL)* structure, since the CAEL framework is most widely used to assess the S&Ls in the real world. Elmer and Borowski (1988) compared the expert system's predictive performance with the logistic regression model. Their results indicated that the expert system has better utility than the traditional statistical techniques. Other studies related to rule-based expert systems in credit scoring domain are: Srinivasan's and Kim's (1988) financial expert system and Srinivasan and Ruparel's (1990) CGX system.

Expert system has a number of advantages. For example, Elmer and Borowski (1988) pointed out that the main advantage of an expert system is its flexibility and potential for default prediction model development. In contrast, expert systems also have some drawbacks. Vedder (1987) argued that the most serious shortcoming is the lack of robustness. This means that an expert system can only be applied to solve a narrow or a specific problem. Qureshi et al. (1998) also listed some drawbacks relative to an expert system, including lack of innovation, difficulty to face changing environment and high development costs.

2.3.2.7 **Artificial Neural Networks**

From the early 1990s, another artificial intelligence or machine learning technique, the *Artificial Neural Networks (ANN)*, was successfully applied to financial distress prediction studies. The most popular ANN algorithm in the financial distress prediction domain is the *Multilayer Perceptron (MLP)*. A MLP has three main components: input layer, hidden layer and output layer. The input layer is responsible for receiving information from the outside environment and transferring it to the hidden layer. In the hidden layer, a neuron will assign a series of weights to the inputs, cope with the information via a training process, and then forward the results with weights to the output layer. The training process can be viewed as a weighting determination process.

The most frequently used algorithm for the training process is the *Back Propagation Algorithm (BPA)*. Thomas et al. (2002) pointed out that BPA first calculates the difference between the expected output value and the observed output value (called *error*) in the output layer. The next step is to distribute the error back to the network in terms of a weight and to adjust the weight to decrease the error. The process is repeated for all cases, called an *epoch*. After several epochs training, the learning error will reduce to a minimum level and the training process ends.

Trigueiros and Taffler (1996) mentioned some advantages of MLP, such as the independence from statistical distribution assumptions and the ability to deal with a wide complex interaction among independent variables. However, MLP has limitations. For example, it does not provide adequate significance tests and requires considerable computer power and skills (Tam and Kiang, 1992). Moreover, neural network cannot easily explain the predictive results conceptually (Sung et al., 1999). Finally, Trigueiros and Taffler (1996) argued that the generalization process of neural network may suffer from overfitting, since enough nodes can be used to find the best fit based on the sample data. Again, cross validation process can be employed to solve potential overfitting problem and it will be illustrated in Chapter Seven.

2.3.2.8 Support Vector Machine

In the late 1990s, another machine learning technique, *Support Vector Machine (SVM)*, was introduced to deal with the classification problem. Fan and Palaniswami (2000) applied SVM to select the financial distress predictors in the Australian market. They pointed out that SVM created an optimal separating hyperplane in the hidden feature space in terms of the principle of structure risk minimization, and used the quadratic programming to obtain an optimal solution. In addition, SVM is able to classify healthy and distressed firms based on some complex data patterns by generating a highly nonlinear separating surface. It is achievable by employing *Kernel* functions (Lee, 2001). SVM has been successfully applied to many disciplines, such as DNA analysis (Brown et al., 2000), breast cancer diagnosis (Lee et al., 2000) and face detection (Osuna et al., 1997).

Fan and Palaniswami (2000) constructed a financial distress prediction model using variables from Altman (1993), Lincoln (1982) and Ohlson's (1980) studies. They then carried out a comparative analysis based on prediction performance in terms of four different credit scoring techniques: *SVM*, *Linear Discriminant Analysis (LDA)*, *Multi-layer Perception (MLP)* and *Learning Vector Quantization (LVQ)*. The results showed that SVM obtained the best results, followed by MLP, LVQ and LDA.

However, Platt (1999) argued that a large number of quadratic programming in SVM training is time consuming and is too complex to implement in the real world, especially in the engineering community. As a result, he introduced a new algorithm, *Sequential Minimal Optimization (SMO)*, to improve the SVM training time. Unlike the previous SVM training methods, SMO does not require the numerical quadratic programming optimization process and any extra matrix storage. Therefore, although SMO requires more iterations to converge, it only requires a few operations in each step and is overall, quick to run (Cristianini and Shawe-Taylor, 2000). It was found that SMO has better performance than other SVM training methods, including better scaling with training sample size.

2.3.2.9 Other Credit Scoring Techniques

Other methodologies were also applied to the financial distress prediction research area and have shown good performance—for instance, *Human Information Processing* (Libby, 1975; Casey Jr., 1980), *Survival Analysis* (Lane et al., 1986; Luoma and Laitinen, 1991), *Mathematical Programming Method* (Gupta et al., 1990), *Rough Set Approach* (Dimitras et al., 1999; Mckee, 2003), *Multidimensional Scaling Approach* (Mar-Molinero and Serrano-Cinca, 2001) and *Genetic Programming Method* (McKee and Lensberg, 2002; Lendberg et al., 2006). The development of the accounting-based (or credit scoring) default prediction model was summarized in Table 2.3 from a historical point of view:

Table 2.3 Development of Accounting Based Default Prediction Model

Time Period	Credit Scoring Technique
Mid 1960s	Univariate Analysis
Late 1960s	Multiple Discriminant Analysis
Mid 1970s	Human Information Processing
Early 1980s	Conditional Probability Model
Mid 1980s	Recursive Partitioning and Survival Analysis
Late 1980s	Expert System and Mathematical Programming
Early 1990s	Artificial Neural Networks
Late 1990s	Rough Set Approach
Early 2000s	Support Vector Machine, Sequential Minimal Optimization, Genetic Programming Method and Multidimensional Scaling Approach

2.4 Key Issues Regarding Financial Distress Prediction Research

2.4.1 Distressed Sample Selection (Database Problem)

Keasey and Watson (1991) pointed out that many studies selected distressed firms from the Moody's Industrial Manual or the Compustat Industrial Files in the US market (DataStream or Extel in the UK market). Usually these databases only include publicly listed companies or large private firms. If a financial distress researcher intends to focus on the private sector or small companies, the representative-ness of these databases is limited. In addition, the neglect of small firms is a severe problem, since small firms are more likely to face financial distress than large companies (Franklin, 1981; Hudson, 1986). As a result, it is worthwhile to explore the default situation of small firms than that of large firms.

2.4.2 Healthy Sample Selection (Paired-Sample Design)

A number of previous studies employed paired-sample design to select healthy firms in order to avoid the influences from the variety of industries and sample size – for example, Beaver (1966), Altman (1968), Wilcox (1973), Norton and Smith (1979) and Zavgren (1985). However, the paired-sample design may cause oversampling bias. Oversampling bias is caused by considering more distressed firms (relative to healthy firms) in the sample than its proportion in whole population (Dietrich, 1984). Zmijewski (1984) examined the oversampling bias by using the probit analysis. Results showed that oversampling bias exists, although these biases can be eliminated through an adjustment procedure.

Another important consideration regarding paired-sample design is that firm size or industry is potentially good predictor for forecasting financial distress (Beaver 1966). Ohlson's (1980) study indicated that the size measure: log (total assets/GNP price-level index) is a statistically significant variable. Drawing on this, paired-sample design may not a good strategy for selecting healthy firms.

2.4.3 Data Collection—Timing Consideration

As mentioned in Section 2.2.3.4, Ohlson (1980) pointed out that the most recent financial statement prior to default for a distressed firm may only be available after bankruptcy is filed—if this company files for bankruptcy between the date of financial statements releasing and the date of the financial year. Ohlson suggested that for these companies, the financial statements prior to the financial distress year should be viewed as the last report, since reports after financial distress would usually include the adjustments from auditors in light of the bankruptcy filing.

Another timing issue is the neglect of new firms. New firms, which are usually small, also have a great likelihood of facing financial distress (Knott and Posen, 2005). Most previous studies evaluated the performance of the default prediction model in the long term, such as five years prior to financial distress. As new firms usually do not have long-term data, new firms tend to be excluded in a financial distress prediction model. This is a major disadvantage of previous financial distress model studies. They only reflect the real default situation to a certain extent.

2.4.4 Variable Selection

A review of past academic literature has indicated that most financial distress prediction models are based on popular quantitative financial ratios–for instance, Altman's (1968) Z-score model and Ohlson's (1980) logistic regression model. However, these models tend to ignore significant qualitative factors, such as a company's execution ability. In fact, many renowned credit-rating companies including Moody's, S&P, and Fitch take into account both quantitative and qualitative factors when carrying out credit assessment but attribute more importance to qualitative rather than quantitative factors in the process (Moody's Investors Service, 1998 and 2002; Fitch Ratings, 2000 and 2001; S&P, 2002 and 2003).

As mentioned in Section 2.2, incorporating both financial and non-financial data in a prediction model would provide a more reliable framework for evaluating corporate performance. Marais et al. (1984) worked on a commercial bank loan classification model where they employed three different categories of independent variables: financial ratios, financial non-ratio variables and non-financial ratios. They found that the non-financial variables could possess as much explanatory power as financial ratios. Following this, a number of default prediction studies also showed that including qualitative or non-financial variables brought about higher accuracy in a financial distress prediction (Peel et al., 1986; Keasey and Watson, 1987; Becchetti and Sierra, 2003; Kuo et al., 2003; Wu, 2004).

Another important issue related to performance measures is industrial influence. Since every industry has its distinctive characteristics, applying the same kind of variables to different industries may lead to an overly general model that overlooks the specific attributes (Williams and Goodman, 1971; Gupta and Huefner, 1972; Bowen et al., 1982; Mensah, 1984). Platt and Platt (1990) plugged industry-related measures into the bankruptcy model and proved that these industry-relative measures could improve the accuracy of the classification model.

Some researchers are interested in whether the general price changes have impacts on historical accounting ratios. Norton and Smith (1979) examined the financial distress prediction power by comparing two models: the general price-adjusted ratios model and historical ratios model. These two models resulted in similar prediction performance. Ketz (1978) also carried out a similar study. He argued that despite similarity in the overall prediction power between the two models, the general price-adjusted ratios model performed better in term of the error rate of misclassified failed firms. Mensah (1983) developed a specific *Price Level Adjusted* model and argued that it is difficult to conclude the utility of the general price-adjusted ratios model, since its performance varies in terms of different modelling techniques.

The external environmental factors will also have great impacts on corporate performance. Liu (2004) examined the relationship between macroeconomic factors and business failures in the UK market from 1966 to 1999 and the results showed that the business failure rates are related to some macroeconomic factors, such as, interest rates, both in the short run and in the long run. Rose et al. (1982) adapted various macroeconomic indicators, such as unemployment rate, into a financial distress prediction model, and they pointed out that these macroeconomic indicators are able to enhance the performance of the prediction model. Furthermore, Mensah (1984) also pointed out that different macroeconomical environments may affect the accuracy of the bankruptcy predictive model. A number of studies applied Altman's (1968) Z-score model to different time periods and results showed a worse performance compared to the Altman's (1968) original results (Moyer, 1977; Begley et al., 1996; Grice and Ingram, 2001). It can be concluded that various macroeconomical factors and environments have significant impact on the performance of default prediction models.

A company's sustainability is also significant, since it measures a company's ability to face its future obligations. Analysis of a company's sustainability must be based on cash flow, rather than on earnings in the accounting statements (for earnings include non-cash items that cannot reflect a company's ability to pay back interests or principal; S&P, 2003). In general, cash flow measures in the default prediction studies can be divided into two groups: 1) income plus depreciation and amortization (or profitability cash flow measures), and 2) income adjusted for all accruals (or operating cash flow measures). Gombola and Ketz (1983) examined the patterns of a series of cash flow measures by conducting a factor analysis and found that the profitability cash flow measures had similar pattern to certain accounting profitability measures. Moreover, the patterns between the profitability cash flow measures and the operating cash flow measures were different.

A number of studies used the profitability cash flow measures as a proxy for cash flow in the bankruptcy prediction studies (Beaver, 1966; Norton and Smith, 1979; Mensah, 1983). Most concluded that these cash flow variables can improve the prediction ability. However, the same cannot be concluded in some other studies. For example, Sharma and Mahajan (1980) argued that the best model did not include any cash flow measure and the classification accuracy rate of the model without cash flow variable is 91.67% one year before financial distress.

With regards to the operating cash flow measures, the prediction ability also presented debatable results in the previous studies. For example, Casey and Bartczak (1985) pointed out that the operating cash flow measures did not improve the default prediction ability and this conclusion is consistent with the results of Casey and Bartczak (1984), Gentry et al. (1985) and Gombola et al. (1987). In contrast, Largay III and Stickney (1980) argued that the operating cash flow measures have better prediction ability than other financial ratios based on the bankruptcy case of W.T. Grant Company. Similarly, Takahashi et al. (1984) showed that the operating cash flow measures could improve the prediction utility of the bankruptcy prediction model in terms of the Japanese market. Finally, Aziz et al. (1988) compared the cash flow based model with both Altman's Z-score model and ZETA model and they indicated that the cash flow based model had favourably performance as both Z-score and ZETA models.

The variability of financial ratios is also an important consideration for variable selection in the corporate default prediction domain. Dambolena and Khoury (1980) employed three variability measures: the standard deviation, standard error and coefficient of variation of the financial ratios into a financial distress prediction model. They showed that the incorporating of variability measures could enhance the prediction performance. This conclusion was confirmed by Betts and Belhoul's (1987) study in the UK market.

From the discussions above, it is evident that a range of variables have been considered in the default prediction domain and some of them have shown to have the ability to improve the prediction performance. However, the primary argument of the financial distress prediction model is that, unlike the performance measurement systems, such as the balanced scorecard, financial distress prediction model lack theoretical groundwork for variable selection. Often, financial distress researchers select independent variables for model construction only based on successful prediction performance or popularity in previous studies. Obviously, such a variable selection method is limited and fails to provide a holistic framework for research in financial distress.

Another limitation of previous studies of variable selection is they have not taken into account subjective assessment. Although one of the criteria for variable selection in Altman's (1968) study is the judgement of the analyst, this criterion is not updated. A fundamental question arises: 'Are the performance measures in the previous default prediction studies important in the real world?' If the selected variables are not vital, then the practical applicability of the default prediction model is questionable. In the present research, adopting a practical viewpoint for variable selection will be considered important in assessing financial distress.

2.4.5 Variable Reduction and Key Variable Selection Techniques

Due to the lack of theoretical groundwork for variable selection, most financial distress prediction studies took into account a large number of variables in order to consider all potential useful variables (Jones, 1987). Due to the danger of overfitting when including a large number of variables, there is a need to reduce the number of variables (Zavgren, 1983). *Stepwise Fitting (SF)* and *Principal Components Analysis (PCA)* were the most two popular methods for variable reduction and selection in the previous bankruptcy prediction studies, such as, *SF* for Gilbert et al. (1990) and Baldwin and Glezen (1992) or *PCA* for Pinches et al. (1973) and Libby (1975).

For example, Pinches et al. (1973) carried out a PCA based on a range of financial ratios, which have been examined to be useful in previous bankruptcy prediction and bond rating studies, in order to develop an empirically based classification of financial ratios. These financial ratios can be divided into seven categories: 1) Return on investment, 2) Capital Intensiveness, 3) Inventory Intensiveness, 4) Financial Leverage, 5) Receivables Intensiveness, 6) Short-term Liquidity, and 7) Cash Position. Taffler (1983) also used PCA to reduce the number of variables. He argued that PCA has some advantages, such as, avoiding potential multicollinearity problems, understanding the data better and facilitating explanation of results.

Moreover, Chen and Shimerda (1981) defined the useful financial ratios in the bankruptcy prediction research. They first found 34 financial ratios, which had been proven to have sound prediction ability in seven previous default prediction studies (e.g. Beaver, 1966; Altman, 1968; Deakin, 1972; Edmister, 1972; Blum, 1974; Libby, 1975; Elam, 1975). They conducted PCA to find useful financial ratios based on the seven categories in Pinches' et al. (1973) study. They then identified ten key variables in the default prediction research (see Table 2.4). One drawback of Chen and Shimerda's (1981) study is that the identified ten useful financial ratios for bankruptcy forecasting were selected based on only a few previous studies. Furthermore, as discussed in the previous section, their study did not have a theoretical framework and may lack insights gain from the context of default.

Table 2.4 Key Variables in Bankruptcy Prediction Research

Factor Category	Key Variables
Return on Investment	1. Net income/sales 2. Net income/common equity
Capital Turnover	3. Working capital/total assets
Financial Leverage	4. Funds flow/total debt 5. Funds flow/current liabilities 6. Long-term debt/current assets 7. Retained earnings/total assets
Cash Position	8. No credit interval 9. Quick flow
Receivable Turnover	10. Quick assets/inventory

Source: Modified from Chen, K.H. and Shimerda, T.A. (1981) An empirical analysis of useful financial ratios, Financial Management, 10, 1, p.58

2.4.6 Type I and Type II Error

The most straightforward way to evaluate the performance of a default prediction model is the classification accuracy rate. This is employed widely in previous studies and the prediction success is defined as the joint minimization of Type I and Type II misclassification errors. Type I error is defined as the error to classify a distressed firm as a healthy firm, while Type II error is defined as the error to classify a healthy firm as a distressed firm. Regarding the importance between the Type I and Type II error, it depends on the users of the financial distress prediction model.

For example, from an investor's perspective, the cost of Type I error is higher than the cost of the Type II error. Type I error may cause an investor to lose the entire investment, while Type II error may only cause an investor to lose the potential dividends or capital gains (Koh, 1992). Altman et al. (1977) estimated the misclassification costs based on a role of the commercial bank loan function and showed that the cost of Type I error is 35 times that of Type II error.

In contrast, from a management point of view, Type II error is more costly than Type I error, since Type II error will damage a company's reputation. Furthermore, this company may lose a number of good investment opportunities, since no lenders will provide financial loans to this company.

2.4.7 Model Performance Comparison

A large number of studies compared the prediction performance of the Artificial Neural Network (ANN) with other classification methods and proved that ANN had better prediction performance than other methods (e.g. Salchenberger et al., 1992; Tam and Kiang, 1992; Coates and Fant, 1993; Wilson and Sharda, 1994; Fernandez and Olmeda, 1995; Zhang et al., 1999; Charitou et al., 2004). In particular, Tam and Kiang (1992) compared the classification ability in terms of five credit scoring techniques: MDA, logistic regression, k nearest neighbour approach, decision tree and artificial neural network. Their results showed that neural network presented a better discriminating ability then other techniques.

However, other studies have not given such clear results, (Altman et al., 1994; Boritz and Kennedy, 1995; McKee and Greenstein, 2000). For example, Boritz and Kennedy (1995) compared the prediction performance in terms of neural network, MDA, logistic regression and probit regression by using the variables from Altman's (1968) and Ohlson's (1980) studies. Their results indicated that it is difficult to conclude which modelling technique has superior predictive ability, since the performance of the modelling technique is affected by different sets of variables. Laitinen and Kankaanpaa (1999) confirmed this. They compared firms' prediction performance in terms of six different credit scoring techniques: linear discriminant analysis, logistic regression, recursive partioning, survival analysis, artificial neural network and human information processing. Their results showed that the performance of different credit scoring techniques varied in terms of different time periods, such as one year, two years and three years prior to financial distress. Overall, the performance of various credit scoring techniques is still open to debate in the financial distress prediction research domain.

2.5 Concluding Remarks

This chapter began by introducing the development of performance measurement systems, such as strategic cost management or the balanced scorecard. These performance measurement systems have been widely adopted for performance assessment and strategic control in the real world and have shown good utility. However, despite providing a sound framework for performance evaluation, these systems still present some limitations. For example, most performance measurement systems concentrate on risk assessment within a general business context, and tend to lack the ability to deal with specific industrial sector. A number of studies argued that different industries have different financial characteristics, (e.g. Williams and Goodman, 1971; Gupta and Huefner, 1972) and applying the same analysis to different industries may lead to irrelevant conclusions.

Drawing on the above, the retail industry is selected to illustrate the workings of this research. Due to structural changes in the retail environment, retail risk evaluation has become increasingly important. Dawson (2000) argued that one of the important retail research issues for the next five years is *Retail Risk Assessment and Evaluation*. Furthermore, he also suggested that one potential retail research direction is to assess retail risk based on a theoretical framework.

Despite providing frameworks for performance evaluation, another limitation is that few empirical studies examine the prediction power of these performance measurement systems. Normally, decision makers are more interested in performance prediction, since they need to design strategies for future operations under the uncertainty surroundings. Credit risk prediction is one of the approaches that can be employed to predict a company's performance. This chapter, therefore, discussed the development of the corporate credit risk prediction, and particularly—the corporate default prediction development.

Corporate financial distress prediction models can be divided into the market based model and accounting based model. The market based model, such as Moody's KMV model, is founded on the Merton's (1974) option pricing framework. By calculating the distance between the market value of assets and book value of liabilities, the likelihood of a company to face financial distress can be obtained. The present research concentrates on the development of the accounting based model, since it has strong connection with performance measurement in the management accounting domain.

Accounting based models employ credit scoring techniques to predict corporate default. From the definition in Thomas et al. (2002), credit scoring can be regarded as a collection of techniques derived from statistics or artificial intelligence domains, and used for individual, corporate or government credit risk assessment. In connection with corporate credit risk evaluation, the primary aim is to establish a classification rule to distinguish the credit risk between 'healthy firms' and 'distressed firms' with the purpose of assisting lenders to make loan decisions. Following a credit scoring process, each company is attributed a credit score based on some specific measures. These credit scores enable the forecasting of a company's probability to face financial distress and hence, achieve the goal of predicting performance.

A number of credit scoring techniques can be utilized to develop corporate financial distress prediction models. In this chapter, several credit scoring techniques: Univariate Analysis, Naïve Bayes, Multiple Discriminant Analysis, Logistic Regression, Recursive Partitioning, Expert Systems, Artificial Neural Network, Support Vector Machine, and Sequential Minimal Optimization are introduced in terms of a historical point of view. Furthermore, the advantages and limitations of each technique are also briefly presented.

Some key issues regarding the development of corporate default prediction research are also presented in this chapter. Due to the constraint of certain databases, small or private companies are likely to be ignored in many default prediction studies. This causes serious problems. For example, small companies are more likely to face financial distress than large companies. This drawback also limits the default prediction research within the private sector.

A number of previous studies adopted paired-sample design to select the healthy firms with which to compare distressed firms in order to eliminate the influences from various industries and sizes. However, paired-sample design may to produce the oversampling bias and disregard the likelihood of the predicting power of size or industry. Timing for data collection is also an important issue. It is possible to collect financial data for distressed firms after the default. This kind of data is usually not comparable with other financial reports, as it includes the adjustments from auditors in light of the bankruptcy filing. Regarding the issues of sample selection and data collection of this research, they will be illustrated in Chapter Six.

With regards to variable selection, a variety of indicators have been considered in previous studies, such as, qualitative variables, non-financial variables, industry adjusted measures, price adjusted measures, macroeconomical variables, variability of financial ratios and sustainability (cash flow) measures. Not all these measures have the ability to improve the prediction performance. The primary argument of the financial distress prediction model is that unlike the performance measurement systems, financial distress prediction model lack theoretical groundwork for variable selection. Often, financial distress researchers select independent variables for model construction only based on the successful prediction performance or popularity in previous studies. They fail to provide a holistic framework for research in financial distress. Moreover, most previous studies selected predictors without the insights from the context. This will also limit the practical applicability of the default prediction model.

Due to the lack of theoretical framework for variable selection, most previous studies took into account a large number of variables in order to consider all potential indicators. However, too many variables in a default prediction model tend to overfit the results, and hence it is necessary to reduce the number of variables and to select the key variables. The most frequently used approaches to reduce and select variables are Stepwise Fitting and Principal Component Analysis. In this thesis, the variable reduction and key variable selection procedure will be presented in Chapter Six.

Although a range of credit scoring techniques can be used to develop financial distress prediction model, the comparison of their prediction ability is still inconclusive. The current study will compare five different credit scoring techniques: Naïve Bayes, Logistic Regression, Recursive Partitioning, Artificial Neural Network, and Sequential Minimal Optimization. Moreover, as different decision makers are interested in different type of errors, the analysis of Type I and Type II errors will also be carried out. The modelling procedure and the assessment of the model's prediction ability and practical applicability will be presented from Chapter Seven to Chapter Eleven.

From the discussions above, it is obvious that a gap between the studies in performance measurement systems and performance prediction models exists. On the one hand, performance measurement systems are founded on a theoretical framework that focuses on a general business context rather than a specific industrial sector. Moreover, decision makers are more interested in performance prediction than performance evaluation, but few empirical studies examine the prediction ability of current performance measurement systems. On the other hand, credit scoring models have the ability to provide a platform for forecasting company performance. However, the main argument of these models is that most previous default prediction models lack theoretical framework and realistic view for variable selection.

Measuring and predicting performance are interrelated issues that should be studied together. A series of research questions arise from the literature review. For example:

- Can we develop a retail performance measurement system based on a theoretical framework?
- Can we construct this framework based on both academic and practical considerations?
- Can we apply this theoretical framework to predict retail corporate default by using credit scoring techniques?
- How can we evaluate this financial distress prediction model?
- Can we apply this default prediction model to the real world?

The aim of this thesis is to fill the gap between the previous studies in both performance measurement systems and performance prediction models by overcoming the relevant limitations. In order to achieve this research objective, a fundamental theory should be selected in advance for model framework construction. In this research, the Resource-Advantage (R-A) Theory of Competition is selected for developing the research framework. R-A theory will be introduced in the next chapter (Chapter Three). In addition, the development of the research framework will be presented in Chapter Four and Five.

Part Three

Research Framework Development

- Chapter Three: ***Resource-Advantage Theory of Competition***
- Chapter Four: ***Research Framework Development***
- Chapter Five: ***Survey Examination of the Research Framework***

Chapter Three looks at the fundamental R-A theory, as well as the basic premises of the theory structure and the implementation process. The main purpose is to provide a blueprint for research framework development. Based on R-A theory, *Chapter Four* then focuses on the construction of the model framework. This involves literature review and interviews with practitioners. In *Chapter Five*, a survey analysis is presented. The survey provides real-world insights regarding the importance of performance measures. The primary purpose is to conduct a robust examination of the model framework. A series of comparative analysis are also carried out based on different countries, retail management functions and retail formats. Finally, a case study will illustrate the application of R-A theory.

Chapter THREE

Resource-Advantage Theory of Competition

3.1 Introduction

As discussed in the previous chapter, most financial distress prediction models can be criticized for their lack of theoretical groundwork for variable selection. Here, the Resource-Advantage (R-A) Theory of Competition is selected for developing the research framework. The fundamental premises of R-A theory are highly related to the real retail competition environment. R-A theory is capable of describing the dynamic process and environment of retail competition and considers corporate competition advantage not only based on a company's internal resources, but also based on the influences from the external environment. Hence, it provides a more complete blueprint for construction of a research framework.

This chapter begins by briefly introducing the background of R-A theory. It then concentrates on discussing the fundamental premises of R-A theory in order to demonstrate why R-A model of competition closely reflects the retail competition environment. Following this, a section is devoted to illustrate how R-A theory can be implemented in the practice. Finally, the key issues addressed in this chapter are summarized and the application of R-A theory to the current research is discussed.

3.2 An Overview of R-A Theory

R-A Theory was developed in the mid-1990s and it has been applied to many different disciplines, such as, marketing, management, economics and ethics (Hunt and Arnett, 2003). The pedigree of R-A theory draws on eleven different research traditions, such as heterogeneous demand theory, resource-based tradition or competence-based tradition. The heterogeneous demand theory contributes to R-A theory in explaining why different market segments exist in the same industry (Hunt, 2000:10). In addition, R-A theory takes into account the resource-based perspective regarding imperfectly mobile resources (Grant, 1991; Sanchez and Heene, 1997); this contributes to R-A theory in explaining why a firm can maintain its market position despite the efforts from its competitors. Furthermore, the competence-based tradition defines competition as a dynamic, disequilibrium process where major drivers, such as learning from competition, keep the dynamism in competition (Teece, et al., 1997; Dickson, 1996). Drawing on the insights from eleven different research traditions, Hunt and Arnett (2003) defined R-A theory as:

> *"an evolutionary, disequilibrium-provoking, process theory of competition, in which innovation and organizational learning are endogenous, firms and consumers have imperfect information, and in which entrepreneurship, institutions, and public policy affect economic performance."*

(Hunt and Arnett, 2003:4)

3.3 The Fundamental Premises of R-A Theory

Hunt and Arnett (2001) pointed out that compared with the traditional perfect competition theory, the foundational premises of R-A theory provide a more realistic description of the competition process. Table 3.1 shows the primary differences between perfect competition theory and R-A theory.

Table 3.1 Foundational Premises of Perfect Competition and Resource-Advantage Theory

	Perfect Competition Theory	Resource-Advantage Theory
P1. Demand is:	Heterogeneous across industries, homogenous within industries, and static.	Heterogeneous across industries, heterogeneous within industries, and dynamic.
P2. Consumer information is:	Perfect and costless	Imperfect and costly
P3. Human motivation is:	Self-interest maximization	Constrained self-interest seeking
P4. The firm's objective is:	Profit maximization	Superior financial performance
P5. The firm's information is:	Perfect and costless	Imperfect and costly
P6. The firm's resources are:	Capital, labour and land	Financial, physical, legal, human, organizational, informational, and relational.

Table 3.1 Foundational Premises of Perfect Competition and Resource-Advantage Theory (Continued)

P7. Resource characteristics are:	Homogeneous and perfectly mobile.	Heterogeneous and imperfectly mobile.
P8. The role of management is:	To determine quantity and implement production function	To recognize, understand, create, select, implement and modify strategies.
P9. Competitive dynamics are:	Equilibrium-seeking, with innovation exogenous	Disequilibrium-provoking, with innovation endogenous.

Source: Hunt (2000) A General Theory of Competition, Thousand Oaks: Sage Publications, p.106

This chapter now turns to the foundational premises of R-A theory and their applicability to the retail competition environment.

3.3.1 Demand

Unlike perfect competition theory, *demand* in R-A theory is heterogeneous not only across industries, but also within an industry. The premise of heterogeneous and dynamic demand in R-A theory helps explain why a range of different formats in the retail industry exists during the same time period. This premise also explains the diversity of a retail company's size, scope and performance. Indeed, McGoldrick (2002) mentioned that changes in the retail environment are becoming more rapid and diverse, given increasingly demanding customers and growing competition between different retail formats and channels. A number of marketing studies concentrate on the exploration of motivations for shopping in order to understand a variety of customer needs, (Tauber, 1972; Buttle, 1992).

3.3.2 Customer Information

R-A theory assumes that customers have imperfect and costly information. This premise implies that customers need to contribute their time and cost to find a product or service to satisfy their needs. Hunt (2000: 112) pointed out that trademarks are important for consumers to reduce searching cost, since they can be viewed as indicators of product quality. This premise can also be applied to retailing: retailers value brand development because they know customers wish to reduce product searching costs (Chang, 2002).

3.3.3 Human Motivation

R-A theory presumes that human motivation is constrained self-interest seeking, since it will be affected by personal moral codes (Hunt, 2000: 113). This is also particularly useful for studies on retailing. For example, the consumption behaviour of 'Ethical Customers' will be affected by their ethical beliefs, such as animal welfare, human rights, or other ethical issues. 'Green Customers' will be affected by their concerns regarding the environmental protection (Varley and Rafiq, 2004). Smith and Cooper-Martin (1997) examined the relationship between product harmfulness and the customer vulnerability. They pointed out that the customer vulnerability increases, as both ethically sound and ethically unsound products are presented. In addition, Babin et al. (2004) also argued that the customer's future shopping intentions were affected by ethical perceptions. These examples support the argument that personal moral codes have considerable impacts on customer consuming behaviour.

3.3.4 Firm's Objective and Information

In the perfect competition environment, the objective of profit maximization is achievable, since a firm's information is perfect and costless. However, in reality, information is always limited and costly, and hence, maximizing profits is not a viable objective. In R-A theory, markets are in disequilibrium whereby seeking superior financial performance can be an appropriate objective. Firms will always take actions to improve their financial performance rather than maximize profits, and this also helps them to achieve other objectives, such as, social responsibility. (Hunt, 2000: 127). The objective of higher financial performance is useful for explaining the dynamism of the retail competition environment (Hasty and Reardon, 1997).

3.3.5 Resources

In perfect competition theory, the primary resources in a company are those quantifiable and easily differentiated elements of production: land, labour and capital. Perfect competition theory tends to ignore a number of significant intangible factors and does not reflect the real world. R-A theory regards intangible factors, such as entrepreneurship or a company's relationship with its customers, as important resources that can enhance a company's market position (Hunt and Arnett, 2003). With regards to corporate internal resources, R-A theory considers both tangible and intangible factors and divides them into seven resources (see Table 3.2).

Table 3.2 Internal Resources

Internal Resources	Examples
Financial Resource	Cash reserves and access to financial markets
Physical Resource	Plant, raw materials, and equipment
Legal Resource	Trademarks and licenses
Human Resource	The skills and knowledge of individual employees and the entrepreneurial skills
Organizational Resource	Controls, routines, cultures, and competences for entrepreneurship
Informational Resource	Knowledge about the market segment, competitors and technology
Relational Resource	Relationships with competitors, suppliers and customer

Source: Modified from Hunt (2000) A General Theory of Competition, Thousand Oaks: Sage p.128

The internal resource classification rule of R-A theory is very close to the resource definition within the retail competition environment. Taking Varley and Rafiq's (2004) grouping of retail resources into physical assets, human resources, financial resources and intangible resources (such as retailer's brand or image), it is obvious that R-A theory is capable of incorporating a large number of internal resources and, hence R-A theory provides a more complete framework for explaining retail competition.

3.3.6 Resource Characteristics

With regards to the resource characteristics, R-A theory assumes that resources are heterogeneous in the same industry. This assumption means that each company has its own distinctive resources. Furthermore, R-A theory also assumes that resources are imperfectly mobile. This implies that resources are not commonly, easily or rapidly to exchange in the market. These resource characteristics provide R-A theory with a platform to explain why firms can retain its competitive advantage in the market regardless of efforts from competitors.

3.3.7 Role of Management

In perfect competition, the role of management is limited. Its main objective is to implement the production function by using the production elements in order to achieve the goal of profit maximization in the short run. In other words, management is viewed as inflexible and lacking strategic vision in the perfect competition environment. In contrast, R-A theory defines the role of management in a strategic manner:

> *"The role of management is to recognize and understand current strategies, create new strategies, select preferred strategies, implement the strategies selected, and modify strategies through time"*
>
> (Hunt, 2000:130)

Hunt (2000:130) further explained each element in terms of this definition. *'Recognize and understand'* means management need to have the ability to know the market demand and accurately identify its market position; *'Create'* implies innovation ability; *'Select'* indicates strategy selection ability; *'Implement'* denotes decision making ability for dealing with various activities over time; *'Modify'* emphasizes the ability to modify or to abandon underperforming strategies through learning (from competition process). Drawing on the above, management in R-A theory plays a more realistic, flexible role emphasizing long-term direction. Moreover, given the highly dynamic nature of retail competition, this premise also provides a more appropriate description of the role of management in retailing.

3.3.8 Competitive Dynamics

Unlike perfect competition theory, R-A theory assumes that competition is an evolutionary and disequilibrium-based process (Hunt, 2000:132). R-A theory adapted viewpoints from evolutionary economics according to which the survival of a firm in a dynamic competition environment is based on genetic variation, selection and retention. Since the competition environment changes over time, the market composition will also vary. Firms are capable of responding to environmental challenges in order to survive in the market (Sanchez and Heene, 2003).

R-A theory also assumes that innovation is endogenous. As mentioned in Section 3.3.4, firms will always take *actions* to seek superior financial performance and contribute to the dynamism of the market (Hunt and Arnett, 2003). Innovations, which include both proactive innovation and reactive innovation, are the primary elements to uphold the dynamism of the market. Proactive innovation, similar to the entrepreneurial spirit, is the motivation to seek superior financial performance (Hunt, 2000:87). Reactive innovation, such as resource imitating or resource creating, is prompted from the competition environment through a learning process (Hunt, 2000:88).

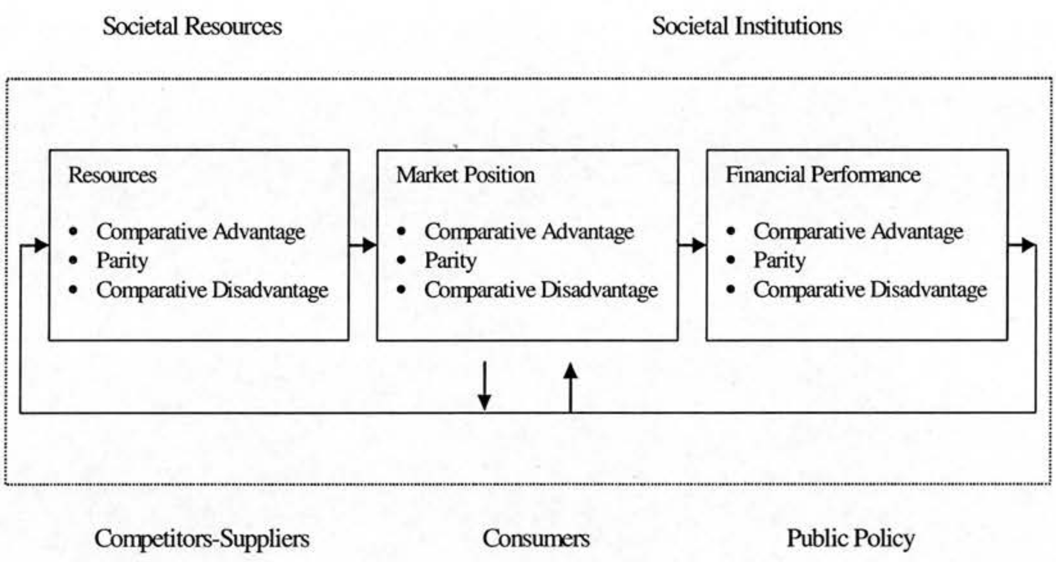
The competitive dynamics underlying R-A theory again reflect the retailing environment. Innovation leads the success of retail revolution, (Dawson, 2001_a). Dawson (2001_b) developed an Innovation-Productivity model for new commerce. He argued that innovation is generated by combining both technology advances and the managerial ability. Moreover, he pointed out that innovation is the major driver for a firm's productivity, as it can lower cost (or lower price) and achieve higher levels of service. As productivity improves, an increase of sales can be expected through the channel control cycle and marketing cycle.

3.4 Implementation of R-A Theory

Having presented the premises of R-A theory, this chapter now turns to discussion its implementation. According to Hunt and Arnett (2003), four main elements form the R-A competition process: market segments, heterogeneous firm resources, a comparative advantage (or disadvantage) in resources, and a comparative advantage (or disadvantage) in the market position.

In international trade theory, nations gain comparative advantages in terms of their heterogeneous and immobile resources. R-A theory adopted this concept to explain the competition process. As customers have various tastes and the resource characteristics are heterogeneous and imperfectly mobile, each firm will have a comparative advantage based on its distinctive resources in the marketplace. These resources can be tangible or intangible. In other words, a comparative advantage in resources will lead a comparative advantage in some particular market segments. Furthermore, as information is limited and costly, a company with comparative advantage in the market will achieve the objective of superior financial performance. The R-A competition process is presented in Figure 3.1:

Figure 3.1 Resource-Advantage Competition Process



Source: Hunt and Morgan (1997) Resource-Advantage Theory: A Snake Swallowing Its Tail of a General Theory of Competition? Journal of Marketing, 61, pp.78

The competitive position matrix can be used to identify a company's market position as shown in Table 3.3. Hunt (2000; 138) pointed out that a company's competitive position can be decided by relative resource costs and relative resource-produced value. If a company can occupy one of three cells (cell 3, cell 2 or cell 6) in the competitive position matrix, it will achieve superior comparative advantage in the marketplace, and achieve the goal of superior financial performance.

Table 3.3 Competitive Position Matrix

		Relative Resource-Produced Value		
		Lower	Parity	Superior
Relative Resource Costs	Lower	1 Indeterminate Position	2 Competitive Advantage	3 Competitive Advantage
	Parity	4 Competitive Disadvantage	5 Parity Position	6 Competitive Advantage
	Higher	7 Competitive Disadvantage	8 Competitive Disadvantage	9 Indeterminate Position

Source: Hunt and Morgan (1997) Resource-Advantage Theory: A Snake Swallowing Its Tail of a General Theory of Competition?, Journal of Marketing, 61, p.78

Looking back at Figure 3.1, it is obvious that the R-A competition process will be affected not only by the company's internal resources, but also by external environmental factors, including societal resources on which firm's draw, societal institutions that structure economic actions, actions of competitors and suppliers, consumer behaviour, and public policy decisions. A range of retail studies suggest that prior to designing a retail strategy, it is necessary to evaluate the influences from both internal and external retail environments (Merrilees and Miller, 1996; Walters and Hanrahan, 2000; McGoldrick, 2002). As a result, the R-A competition process provides a more complete and appropriate blueprint for research framework construction.

Finally, management plays a more flexible role and emphasises strategic vision in the long-term within the R-A competition theory. Through innovation and learning, management is continuously designing strategy for the purpose of seeking superior financial performance. This process creates a competitive environment that is dynamic, evolutionary and in disequilibrium. The R-A competition process is presented in Figure 3.1 as the competition cycle.

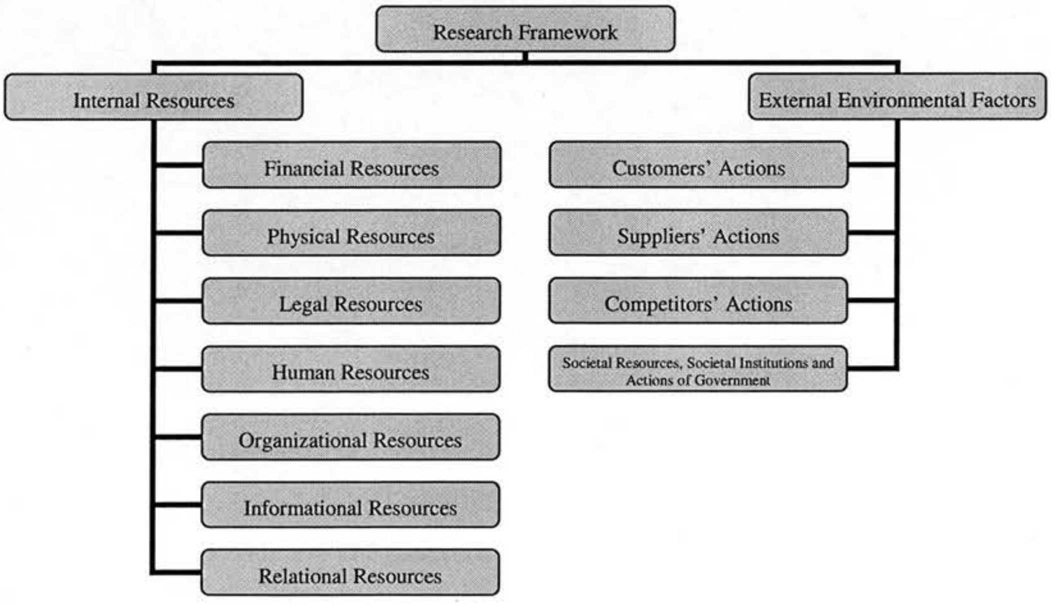
3.5 Concluding Remarks

From the various discussions in this chapter, it is argued that R-A theory provides a more realistic foundation for competition analysis than the traditional perfect competition theory. The applicability of R-A theory to the retail competition environment is high. The R-A theory competition process is also representative of the dynamism in the retailing environment. In summary, it is possible to say that R-A theory is an appropriate and complete theory to explain retail competition process because:

- The fundamental premises of R-A theory are closer to the reality and they are highly related to the retail environment.
- R-A theory can be used to explain the dynamic, evolutionary and disequilibrium retail competition process.
- R-A theory considers a larger range of internal resources than other retail competition theories. Moreover, R-A theory also takes into account external environmental influences as many traditional retail studies.

Drawing on the above, this research will use R-A theory as the fundamental theory for developing research framework. The blueprint of the research framework is simply described in Figure 3.2. In the Chapter Four and Five, more details regarding the framework construction will be introduced.

Figure 3.2 Blueprint of Research Framework



Chapter FOUR

Research Framework Development

4.1 Introduction

Chapter Three presented R-A theory as an appropriate theory to explain the retail competition process. In addition, it was argued that R-A theory provides a blueprint of the research framework in terms of both internal resources and external factors. This chapter will focus on the development of the research framework based on the R-A theory of competition. The main research question in this chapter is:

> *'What is the appropriate set of performance measures for evaluating retail performance?'*

In order to answer this question, a range of retail performance measures will first be collected from secondary sources, such as available literatures or published information from credit-scoring agencies. Second, in an attempt to overcome the limitation of previous default prediction studies in lacking social actor viewpoints for variable selection, this research will incorporate practitioners' opinions for identifying appropriate retail performance measures.

The next section describes methodologies for finding performance measures. It is followed by a literature review of performance measures specific to the retail industry (Section 4.3). Section 4.4 presents an illustration of a pilot study and primary materials obtained from practitioners. The similarities and differences among the viewpoints from different stakeholders will also be discussed. Section 4.5 provides a discussion on framework construction and Section 4.6 presents a summary of findings from the previous sections.

4.2 The Methodology of Performance Measures Investigation

Performance measures will be investigated based on R-A theory by using two types of sources: secondary materials from existing literature and primary materials based on practitioner viewpoints. The secondary materials' review of performance measures is built on reports from credit-rating companies, academic literature and retail management textbooks. Despite finding possible variables for measuring a retail company's performance from secondary sources, it is highly possible that important factors are omitted. Interviewing those stakeholders whose routine job is to evaluate a retail company's performance was considered crucial to enrich the information found through primary sources.

4.2.1 Interview Sampling Strategy

Sampling for interviews is based on the purposive sampling strategy (Gilbert, 2001) or non-probability sampling strategy (Bryman, 2001). The quota sampling approach and snowball sampling approach were selected to be the key sampling techniques. The choice of probability or purposive sampling strategy depends on the aim of the research. Indeed, Gilbert (2001) argued that probability sampling is appropriate if researchers plan to estimate the characteristics of populations or to test an empirical hypothesis. However, if researchers are interested in exploring or developing theory, then purposive sampling strategy is more suitable. As the objective of interview is to explore the insights from practitioners for research framework construction, purposive sampling was deemed an appropriate primary sampling strategy.

Bryman (2001) pointed out that the quota sampling method classifies the population into several different categories based on specific features of the potential interviewees. However, unlike the stratified sampling method, the final sample selection is not carried out randomly. In this research, the criterion for final sample selection is based on the snowball sampling technique—that is, by leveraging the networks of key interviewees.

The total population under study consists of stakeholders whose routine job is to evaluate a retail company's performance. In practice, three interest groups emerge: 1) the management team in retail firms, 2) lenders, and 3) investors. These stakeholders often face the need to evaluate a retail company's performance as part of their routine job in order to make decisions for future operating strategy, long-term loan or investment. The existence of these interest groups called for a quota sampling approach. Hence, the whole population was classified into three groups: retail management, bank managers in business loan departments and industrial analysts in investment institutions. The interviews started in a leading retail company in Taiwan and snowballed into other retail companies, bank loan departments and investment institutions. The final sample selection in each category is based on the accessibility of the interviewees. In total, 25 interviewees were selected and the composition of the interviewees can be described in Table 4.1:

Table 4.1 Composition of Interviewees

Category	Number	Composition
Retail Management	13	The retail companies include: convenience stores, department stores, gas stations, pharmacies, coffee shops, supermarkets, hypermarkets, e-business retailers and other retailers, in Taiwan, UK, Philippine, Japan, China and US.
Bank Loan Managers	8	The bankers include: six Taiwan banks, one US bank and one UK bank
Industrial Analysts	4	The investment institutions include: three Taiwan institutions and one US institution.

4.2.2 Interview Design

A pilot study was carried out in order to ensure the quality of future formal interviews. The participant is previously a project manager in the international department of a leading retail company in Taiwan. The interview is made face-to-face with open questions, and recorded. Strengths and weaknesses of the pilot study are presented in Section 4.4 based on reflections on the literature regarding how to conduct an interview.

During the formal interviewing process, 24 semi-structured phone interviews and one semi-structured face-to-face interview were carried out. The interview period was between end of June and end of August, 2004. Each interview took 20 to 30 minutes. The semi-structured interview design was selected in order to ask key questions relating to the research. At the same time, more questions could be asked based on the replies from the interviewees (Bryman, 2001).

Although the interview questions were slightly different among the groups in this research, the purposes of the interview were the same. For example, with the aim to understand the most important performance factors in the retail industry, researcher usually asked the management a question: 'What are the key factors leading to the success of a retail company?' For the bank managers in the business loan department, researcher would change the question to 'Before you make a loan decision to a retail company, what are the most important performance factors that you consider?' For the industrial analysts in the investment institutions, the question would become: 'What are the most important factors relative to your investment decision to a retail company?' Obviously, these three questions have the same objective: to find out the most important performance factors in the retail industry.

Furthermore, the purpose of investigating and collecting performance measures is not only to understand what the important factors for measuring a retail company's performance are, but also to know how to measure each factor. Thus, each participant was asked: 'How do you measure the impacts of these factors?' or 'What kind of measures will you choose in order to assess these impacts?' As a result, these interviews enhanced the findings from the review of secondary materials, and provided the possibility to construct a more complete framework for developing a performance measurement model in the retail industry.

4.3 Literature Review: Performance Measures in the Retail Industry

This section describes performance factors collected from available literatures, retail management textbooks as well as reports issued by credit rating companies.

Based on the R-A theory, these performance measures can be divided into two categories: internal resources and external environmental factors.

4.3.1 Internal Resources: Financial Resources

4.3.1.1 Profitability

Profitability is very important not only to the retail industry, but also to every industry. S&P (2003) pointed out that a company with higher profitability will have better ability to generate capital internally, attract external capital and fight business adversity. In addition, Gibson and Frishkoff (1983) also argued that increasing profits could cause a rise in market price leading to capital gains.

In the retail industry, different retail sectors encompass significant variations in margin. For example, Fitch Ratings (2000) pointed out that in terms of the ratio of **E**arnings **B**efore **I**nterest, **T**axes, **D**epreciation, and **A**mortization (EBITDA) as percentage of net sales, there are four different margin levels in the European retail industry. These four margin levels are illustrated in Table 4.2:

Table 4.2 Margin Levels in European Retail Industry

Margin Level	Representative Sectors	EBITDA as % of net sales
High margin level	• International branded fashion retailers • Luxury goods retailers	15% ~ 18%
Medium margin level	• Brand-led multiple clothing retailers • "Destination" department stores • Leading specialist non-food retailers	12% ~ 13% (UK) or 7% ~ 10% (Continental Europe)
Low margin level	• Most profitable volume food retailers • Mainstream multiple clothing retailers • Department store chains • Price-led non-food specialists	5% ~ 8%
Weak margin level	• Less profitable food retailers • Discounters • Wholesale Distributions	2% ~ 4%

Source: Modified from Fitch Ratings (2000), Assigning Credit Ratings to European Retailers, p.6

Obviously, the international branded fashion and luxury goods retailers have the highest EBITDA margin in the European area. The aim of enhancing the profitability is usually a primary objective of a retailer and it is achievable by a range of strategies in the retail industry. For example, a number of retailers provide financial services, such as *'Bill Payment Services'*, allowing customers to pay for their public service bills. Stores then earn commissions from this financial service. Furthermore, since the bills collected are also temporarily saved in the bank account, retailers can also gain interest revenue from the financial service. After reviewing various published materials, retail profitability measures can be listed in Table 4.3:

Table 4.3 Key Measures in Evaluating Profitability

Fitch Ratings (2000), Assigning Credit Ratings to European Retailers
• EBITDA as percentage of net sales • EBITDAR as percentage of net sales • EBIT as percentage of net sales • Pre-tax profit as percentage of net sales • Net profit as percentage of net sales • Pre-tax profit / capital employed • Post-tax profit / net operating assets
Moody's (2002), Moody's Approach to Assessing Key Credit Issues in Retailing
• Gross profit margin • SG&A (Selling, general and administrative expense) as percentage of Net sales
S&P (2003), Corporate Ratings Criteria
• EBIT on capital • Operating income as percentage of net sales • Earnings on business segments
TRC (2004), TRC Rating Criteria Corporate Ratings Methodology
• Return of total assets • Operating margins • Earnings by business segment
Gibson and Frishkoff (1983), Financial Statement Analysis: Using Financial Accounting Information
• Net profit margin • Return on total assets • Operating income margin • Return on total equity • Return on common equity • Gross profit margin

Table 4.3 Key Measures in Evaluating Profitability (Continued.)

Ross, Westerfield and Jaffe (1999), Corporate Finance
• Net profit margin • Gross profit margin • Return on total assets • Return on total equity • Dividend payout ratio
Merrilees and Miller (1996), Retailing Management: A Best Practice Approach
• EBIT • Gross profit margin • Net profit margin • Cost of sale percentage • Return on equity
Hasty and Reardon (1997), Retail Management
• Return on equity • Return on net sales • Return on total assets • Net profit margin

4.3.1.2 Liquidity

Liquidity measures the ability of a company to face its short-term obligations (Lev, 1974). A company with a good liquidity usually has better ability to transfer its assets into cash as compared with companies with low liquidity. However, high liquidity is not always good, since it also implies low return on investment (Ross et al., 1999). The most common three liquidity ratios are current ratio, quick ratio, and cash ratio (Foster, 1978; Zimmerman et al., 1990)

4.3.1.3 Sustainability

As with liquidity, sustainability also measures a company's ability to face its future payments. However, the difference between liquidity and sustainability is that sustainability is based on the cash flow framework. Fitch Ratings (2000) argued that if a company has good sustainability, it would have the ability to combat inflation and deliver earnings growth to shareholders. S&P (2003) also mentioned that an analysis of a company's sustainability must be based on cash flow, rather than on earnings in the accounting statements. Accounting earnings usually include non-cash items, such as depreciations, and cannot reflect a company's ability to pay back interests or principal. Furthermore, McGurr and DeVaney (1998) also pointed out that the cash flow based measures have an impact on the accuracy of a retail default prediction model. Drawing on above, the adequacy of cash flow is significant to assess the ability of a retail company to face its principal payments.

As mentioned in Chapter Two, cash flow measures can be divided into two groups: 1) income plus depreciation and amortization (or profitability cash flow measures), and 2) income adjusted for all accruals (or operating cash flow measures). With regards to profitability cash flow measures, the most common measures are: interest cover, dividend cover and fixed charge cover. Fitch Ratings (2000) argued that the most important retail profitability cash flow ratio is the fixed charge cover, since lease finance is very common in the retail industry and the cost of operating lease charges routinely far exceeds the cost of unrestricted-use financial debt. Moreover, Fitch Ratings (2000) also pointed out that the interest cover and the fixed charge cover should be greater than one, as retail companies can deliver moderate accelerated growth while still remaining cash-generative on a net basis. Regarding the operating cash flow measures, they are presented in Table 4.4:

Table 4.4 Key Measures in Evaluating Sustainability

Fitch Ratings (2000), Assigning Credit Ratings to European Retailers
• Net operating cash flow / Gross Capex • Net operating cash flow / Maintenance Capex • Cash dividend cover ((NOCF + cash dividends) / cash dividends)
S&P (2003), Corporate Ratings Criteria
• Funds (working capital) from operations / total debt (adjusted for off-balance-sheet liabilities) • Earnings before interest, taxes, depreciation and amortization (EBITDA) / interest • (Free operating cash flow + interest) / interest • (Free operating cash flow + interest) / (interest + annual principal repayments) • Total debt / discretionary cash flow • Funds (working capital) from operations / capital spending requirements • Capital expenditures / capital maintenance
TRC (2004), TRC Rating Criteria Corporate Ratings Methodology
• Funds from operations to total debt • Funds from operations to capital spending requirements • Free operating cash flow to total debt

4.3.1.4 Leverage

The degree of leverage can affect the ability of a retail company to face its long-term obligations. Distinguishing financial and operating leverage is very important in the retail credit analysis. Many retailers with large physical store work display very high operating leverage. However, high operating leverage is not risky in this situation, since it is caused by consumer demand (Fitch Ratings, 2000). Hence, it is possible to focus on financial leverage analysis (Fitch Ratings, 2000).

Total debt to EBITDA is the ability of a company's cash flow to cover long-term debt. According to Fitch Ratings (2000), if a retailer's Debt on EBITBA ratio is less than one, the retailer has slight leverage; if the ratio is between 1.0 and 2.5, the retailer is moderately leveraged; if it is over 2.5, the retailer is heavily leveraged; if it is over 3.0, the retailer will not be rated as an investment grade.

Regarding the asset cover analysis, the most commonly used measure is debt ratio, which is total debt divided by total asset. According to Fitch Ratings (2000), a debt ratio of less than 50% is regarded as moderate leveraging, while a debt ratio of more than 100% is high leveraging in the retail industry. Fitch Ratings (2000) also considered the influences from market value and the measure of net debt divided by market capitalization is also a significant variable for evaluating retail performance. Table 4.5 is an arrangement of leverage measures from various sources.

Table 4.5 Key Measures in Evaluating Leverage

Fitch Ratings (2000), Assigning Credit Ratings to European Retailers
• Debt / EBITDA • Leased-adjusted net debt / EBITDAR • Debt ratio • Net debt / market capitalization
S&P (2003), Corporate Ratings Criteria
• Gearing ratio • Total debt / (total debt + market value of equity)
Ross, Westerfield and Jaffe (1999), Corporate Finance
• Gearing ratio • Debt to equity ratio • Equity multiplier

Table 4.5 Key Measures in Evaluating Leverage (Continued.)

Hasty and Reardon (1997), Retail Management
• Total asset to total equity • Net sales to net working capital • Total debt to total equity

4.3.1.5 Activity

Activity indicates the ability of a company to control its assets, debts and equity. For example, if researchers want to measure the goods unsaleable risk, they usually use inventory turnover to measure the risk. In addition, they can use receivables turnover to evaluate bad debt risk. If the receivables turnover is high, the implication is that the company's bad debt risk is lower. Therefore, activity measures are also very significant performance measures in the retail industry. The activity measures are summarized in Table 4.6:

Table 4.6 Key Measures in Evaluating Activity

Moody's (2002), Moody's Approach to Assessing Key Credit Issues in Retailing
• Inventory turnover • Receivable turnover
Gibson and Frishkoff (1983), Financial Statement Analysis: Using Financial Accounting Information
• Total asset turnover • Fixed asset turnover • Operating asset turnover • Receivable turnover • Inventory turnover

4.3.1.6 Financial Scale

Since one characteristic of the retail industry is low-margin, scale is more important in retail than in other industries (Fitch Ratings, 2000). Large companies usually have a number of advantages. For example, they have better risk endurance when they face changes in the economic situation (S&P, 2003). In addition, large firms also have better financial flexibility than small companies. A company with large sales or assets can more easily ask for a loan from a financial institution than a small company. Furthermore, companies leading in the retail industry have better sales generation ability than their smaller counterparts and enable to spread fixed costs (Moody's Investor Service, 1999, 2002). As a result, scale is a very important factor in measuring a retail company's market position.

Although scale can bring a number of advantages, Dawson (2000) argued that the *'bigness'* also has some future challenges. For example, a large retailer company may decrease the ability to retain customer's responsiveness, since its organization becomes more diffuse. Furthermore, as an increase of size or diversity implies an increase of potential competitors, a large retailer may also lose its focus on competition. The Japan Bond Research Institute (JBRI) (2002) created a company performance measurement system by using multivariate statistical methodology. The system is called *Corporate Appraisal System by Multivariate Statistical Analysis (CASMA).* In this system, size measures are also significant variables and they are illustrated in Table 4.7. Fitch Ratings (2000) also suggested three important measures to evaluate retail scale. They are expressed in Table 4.8.

Table 4.7 Size and Relative Financial Measures of CASMA

Key Assessment Factors	Relative Financial Measures
Size	• Total capital employed (Yen, Millions, Logarithm)
	• Number of payrolls (Logarithm)
	• Operation cash flow (Yen, Millions)
	• Ordinary income (Yen, Millions)

Source: Nihon Keizai Shimbun, Inc (2002), Japan Well Performance Companies Survey, Online. Available at (In Japanese): http://www.nikkei.co.jp/report/02casma1.html

Table 4.8 Key Measures in Evaluating Scale

• Net sales
• Gross operating margin
• Market share by retail sector

Source: Fitch Ratings (2000), 'Assigning Credit Ratings to European Retails', p. 4

4.3.2 Internal Resources: Physical Resource

A retailer's store location can be regarded as a physical resource. Indeed, Fitch Ratings (2000) pointed out that the most important factor to enhance a retailer's

'Reach Ability' (that is, the ability to serve customers within a certain geographic area) is store location. Reach can be viewed as the ability of a retailer to catch customers. Reach ability is also a measure of competitiveness, since it can increase sales growth, especially under a low-inflation environment. Customers usually have greater purchasing power under a low-inflation environment than in the high-inflation environment. Thus, under these circumstances, a low-inflation environment may encourage customers to spend more, allowing a retail company with better reach ability to increase sales.

Fitch Ratings (2000) mentioned several methods to improve the reach ability. First, reach ability can be improved by diversifying store format, since sales will increase by catering to different consumer's shopping missions or different times of day. Second, reach can also be improved by changing the store layouts and product ranges or by adding facilities, such as free parking. Third, the development of alternative channels, such as internet retailing or catalogue mail order, can also enhance a retailer's reach ability. Many leading retail companies have now developed their own e-commerce and have already produced substantial capital gains.

Reach ability can be measured by a retailer's store network including geographic disposition, the population density of catchment areas and the footfall of major outlets. The most important consideration is location selection. Hasty and Reardon (1997) mentioned an old saying: 'the value of real estate is determined by three things: location, location and location.' Many different factors, such as customer's need of convenience, together lead to the development of retail locations. McGlodrick (2002) also pointed out that even if very small differences exist between physical locations, it can cause serious impacts on the accessibility of store and attractiveness to customers. Furthermore, Merrilees and Miller (1994) argued that the top three important retail mix component are customer service, quality merchandise and location. Drawing on the above, location plays a key role in measuring a retailer's performance. Fitch Ratings (2000) pointed out five important measures that can be used to assess a retailer's reach ability, which are presented in Table 4.9.

Table 4.9 Key Measures in Evaluating Reach Ability

- Store numbers
- Trading area
- Distribution of sales by format and channel
- Size of catchment area in population terms
- The footfalls of major outlets

Source: Fitch Ratings (2000), Assigning Credit Ratings to European Retails, p. 4

4.3.3 Internal Resources: Legal Resources

Brand image or trademark can be viewed as a legal resource, since it is protected by law and competitors cannot take advantages of it. Aaker (1991, 1992) argued that the brand value can benefit both customers and retailers. For example, customers can be more confident to make a patronage decision and reduce the product searching costs. Furthermore, for retailers, brand strength can increase the customer loyalty and lead the marketing strategy more efficiency.

Fitch Ratings (2000) also pointed out that brand strength has many advantages such as, reducing advertising costs (Dawson, 1995), allowing higher margin on private label ranges, increasing negotiating power with institutional landlords and providing protection from economic downturn. Brand strength can be created by improving store environment, product range and customer service (Moody's Investor Service, 1999). Fitch Ratings (2000) assessed brand strength by using several important measures. These measures are expressed in Table 4.10:

Table 4.10 Key Measures in Evaluating Brand Strength

- Advertising expenses as percentage of sales
- Trends in sales conversion rate
- Market capitalization / net assets
- Frequency of store remodeling
- Frequency of marketing strategy redirecting

Source: Fitch Ratings (2000), Assigning Credit Ratings to European Retails, p. 11

4.3.4 Internal Resources: Human Resources

Walter and Hanrahan (2000) pointed out that since customer satisfaction is the key factor for ensuring sales and profits, customer service and the staff response to customers have become major concerns in the retail industry. If quality of service provided by staff is low, customers may leave with a negative shopping experience. Rudolph et al. (2000) pointed out twelve causes of negative shopping experiences and five of these were due to staff. As a result, retailers always attempt to increase the quality of their staff through a sound education and training system.

Another serious problem regarding human resource management in the retail industry is high staff turnover rate (Marchington, 1994). Moody's Investor Service (1999) mentioned two important measures to evaluate the ability of human resource management: how to maintain knowledgeable staff and how to maintain appropriate customer service. This shows that reducing staff turnover rate is an important task in the retail society. Merrilees and Miller (1996) also suggested some important variables for evaluating the performance of human resource management, as listed in Table 4.11:

Table 4.11 Key Measures in Evaluating Human Resource Management

- Internal customer satisfaction, based on quality, timeliness and responsiveness
- Job satisfaction
- Turnover
- Staff orientation and training
- Communication
- Performance feedback
- Training effectiveness
- Absenteeism
- Staff grievances
- EEO/AA complaints

Source: Merrilees and Miller (1996), Retailing management, a best practice approach, p. 405 ~ 406

4.3.5 Internal Resources: Organizational Resources

A number of factors can be employed to describe the organization resources in a retail company, such as execution ability, growth power, productivity and diversification. This section will summarise some important organization resources.

4.3.5.1 Execution Ability

Moody's Investor Service (1999) pointed out that the most important retailing performance measurement factor is the relative level of execution capability. Hasty and Reardon (1997) also argued that the five major dimensions of a retail strategy – location, merchandise, price, service and communications—must be supported by good management ability, such as, store operations, logistics, purchasing, marketing, finance and technology, in order to achieve the objective of high service quality.

Most previous literatures evaluate the execution ability of a retail company by management functions. For example, Moody's Investor Service (1999) focused on the merchandising, supply chain and technology; Merrilees and Miller (1996) listed some non-financial organizational performance indicators in terms of supply, merchandising, support services, selling and marketing; Fitch Ratings (2000) considered merchandising, buying, technology and logistics. These measures to evaluate each management function are presented in Table 4.12.

Table 4.12 Key Measures in Evaluating Execution Ability

Moody's Investors Service (1999), Moody's Evaluates Key Credit Issues in Retailing	
Management Function	Measures
Merchandising	• Define and understand target customers • Selecting appropriate products • Good merchandise assortments • Remodel policy
Supply Chain	• The relationship with suppliers • Logistic ability
Technology	• The ability to generate inventory and operating information
Fitch Ratings (2000), 'Assigning Credit Ratings to European Retails'	
Management Function	Measures
Merchandising	• Product rages • Store layout • Pricing policy
Buying	• The relationship with suppliers • Good global reach
Technology	• The ability to generate inventory and operating information
Logistics	• Efficient warehousing and distribution

Table 4.12 Key Measures in Evaluating Execution Ability (Continued.)

Merrilees and Miller (1996), Retailing management, a best practice approach	
Management Function	Measures
Supply	• Timeliness • Accuracy
Merchandise	• Stock turnover • Stock requested but unavailable
Support Services	• In terms of accounting, marketing, public relations, facilities management and Information system.
Selling	• Complaints • Goods returned
Marketing	• Timeliness • Understanding customer needs • Understanding customer satisfaction

4.3.5.2 Growth Power

Achieving sustainable growth in existing or new markets is a very important performance measure in the retail industry (Moody's Investor Service, 1999, 2002). As mentioned previously, the retail industry is characterized by low-margin. Thus, sales growth becomes the primary driver of earnings in the retail industry. Sales density improvement and expansion in retail trading area can lead to sales growth (Fitch Ratings, 2000).

Fitch Ratings (2000) pointed out that sales density can be improved by increasing footfall (customer visits), sales conversion rate (spending visits/total visits) and average spend-per-visit rate. Footfall is a function of location and it is difficult to predict. Nevertheless, it can be improved by good advertising or other marketing strategies. Sales conversion rate is only important in some retail formats, such as fashion clothing retailer. For grocers, the sales conversion rate is usually over 90% and hence, it is not a significant factor for the grocery sector. Overall, it is mainly for retailers generate sales growth by increasing spend-per-visit rate. Spend-per-visit rate can be improved by changing the product mix.

Expansion is also an effective way to increase sales growth. It can be achieved by opening new stores, create alternative channels and acquisitions. However, expansion will also increase costs, such as rent and occupancy costs. Moreover, expansion may erode sales from current stores and cause negative impacts on sales growth. Finally, external environmental influences should be considered when retailers considering an expansion strategy.

For example, Dawson and Larke (2004) pointed out that as the Japanese economy suffered a recession period in 1990s, a number of Japanese large retailers expanded their business by opening more stores in order to generate sales. However, the results indicated that the expansion presented a lower productivity, high levels of debt and low levels of innovation. Hence, S&P (2003) argued that growth power should be based on stability, since there is very likely risk of over-ambitiousness. A solid expansion plan is the key to increasing sales growth. Many measures can be used to assess a retail company's growth power. Some growth power measures are illustrated in Table 4.13.

Table 4.13 Key Measures in Evaluating Growth Power

Nihon Keizai Shimbun Inc (2002), Japan Well Performance Companies Survey	
Key Assessment Factors	Relative Measures
Growth Power	• Growth rate of total capital employed • Growth rate of number of payrolls • Growth rate of EBIT • Growth rate of equity holders capital
Fitch Ratings (2000), Assigning Credit Ratings to European Retails	
Growth Power	• Like-for like sales growth • Retail market value growth by segment • Store opening program

4.3.5.3 Productivity

Fitch Ratings (2000) evaluated productivity of a retail company in terms of four considerations: cost-based consideration, sales-based consideration, employee-based consideration and cash conversion cycle. Cost-based considerations focus on how retail companies curtail their fixed costs. The main considerations are illustrated in Table 4.14.

Table 4.14 Cost-based Main Considerations

• Establish policy in order to save in energy and communication overheads
• Rising the part-time staffs

Source: Fitch Ratings (2000), Assigning Credit Ratings to European Retails, p. 7

Sales-based considerations concentrate on sales density, which is the average weekly sales per square meter (Fitch Ratings, 2000; Moody's Investor Service, 2002). In order to assess sales density, Fitch Ratings (2000) considered the operating profile of individual retailers, such as format variety, store location and product mix. The average sales density in the Europe retail industry is usually above EUR 150/m^2. Therefore, Fitch regards sales density below EUR 100/m^2 as weak and below EUR 75/m^2 as poor.

Employee-based considerations include measures of profit or sales per employee (Fitch Ratings, 2000). Regarding the financial aspect of productivity, the cash conversion cycle can be utilized to measure the performance of cash management. The cash conversion cycle is the time period between when cash is paid out and when cash is received. Therefore, the shorter the cash conversion cycle is, the better the cash management ability of a company is.

4.3.5.4 Diversification

Diversification can reduce business risk. It can be achieved in terms of product, location, format or customer. For example, a retailer with many products across numerous categories will face less risk than a retailer with a narrow focus. Regarding the geographic diversification, expansion should make strategic sense. For example, Moody's confirmed Wal-Mart's Aa2 ratings, when it acquired Asda in 1999. This acquisition provided Wal-Mart with a platform to enter the UK retail market and should be viewed as a positive impact.

However, Moody's Investor Service (1999) argued that if diversification fails to maintain a target customer group or lack of effective execution, then diversification should be viewed as negative. Due to the rapid development of e-business, a range of retailers have operated a combination of real channel and e-business. Capital expenditures in the Internet channel are a trend. (Moody's Investor Service, 2002) The main measures related to diversification are expressed in Table 4.15:

Table 4.15 Key Measures in Evaluating Diversification

- Capital expenditures in Internet channel
- Diversification impacts
- Whether continue to maintain target customer group?
- Whether lack of effective execution?
- Whether lack of strategic sense?

Source: Moody's Investors Service (1999), Moody's Evaluates Key Credit Issues in Retailing, p. 10; Moody's Investors Service (2002), Moody's Approach to Assessing Key Credit Issues in Retailing, p. 10

4.3.6 Internal Resources: Informational Resources

Hunt (2000) pointed out the informational resources are related to the knowledge about market segment and competitors. Therefore, a good retailer should have the ability to manage various market demands with a sound strategic vision.

4.3.6.1 Market Segment Risk Management

Market segment risk management is significant in the retail industry. According to Moody's Investor Service (1999, 2002) various environmental risks in different market segments will affect the stability of a retail company's future cash flow (Fitch Ratings, 2000, 2001; Moody's Investor Service, 1999, 2002). Examples are fashion risk, seasonality risk and cyclicality risk.

It is difficult to predict fashion change. Many products, such as, toys, apparel or personal electronics can easily become unfashionable with customers. Moreover, retailers usually need to order their inventory far in advance of the selling season. If they cannot predict future fashion trends accurately, their inventory will become out of date and unsaleable. This kind of extreme fashion risk is called *obsolescence product risk*. Therefore, apparel, personal electronics and toy companies face higher fashion risk than other retail format companies. Seasonality will also increase the probability that a retailer cannot face its future principal payments. For example, a toy retailer needs to increase working capital before an important holiday season. If the decisions made on inventory, pricing, or other management actions are wrong, the company may easily face financial distress.

Some retailers, such as durable goods retailer, face higher cyclicality risks, since the time period of inventory turnover is long. For example, the revenue of a furniture retailer depends on the housing turnover situation. If the housing turnover is low, the furniture retailer will not have good performance. The relative business risk in different retailer formats is presented in Figure 4.1. How can we measure these business risks? Fitch Ratings (2000) suggested three important measures as illustrated in Table 4.16.

Figure 4.1 Relative Business Risk in Different Retailer Formats

	Food Prescription Drugs	Auto Parts	Toys Apparels	
Low	———————————	———————————	———————————→	High
	Value Chains Supermarkets	Department Stores Drugstores	Specialty Retailers	

Source: Moody's (2002), Moody's approach to assessing key credit issues in retailing, p. 4

Table 4.16 Key Measures in Evaluating Stability

- Main Market Sales or Profits as Percentage Total Sales or Profits
- Monthly or Quarterly Distribution of Sales and Profits
- Peak Net Debt / Average Net Debt

Source: Fitch Ratings (2000), Assigning Credit Ratings to European Retails, p. 5

4.3.6.2 Strategic Vision

Since the retailing environment changes very rapidly, having strategic vision to forecast and react to market change is a significant factor in evaluating retail performance. Indeed, Moody's Investor Service (1999) argued that a retail company with the ability to adapt quickly and remain flexible would obtain higher ranking. This can also be concluded by Fitch Ratings (2000). Fitch Ratings (2000) pointed out that openness to criticism, willingness to experiment and reaction to initiatives from competitors are key factors to assessing a retail management team's strategic vision. As a result, a sound strategic vision retailer has better ability to obtain market information and to face various market demands.

4.3.7 Internal Resources: Relational Resources

Relational resources are more important in the retail industry than other industries, since retailer directly deal with their customers. Recently, growing concern on the value of the relational marketing has caused many to question the traditional view of marketing (McGoldrick, 2002). Hunt (1997; 431) defined the relational marketing as *'firms are competing through developing long-term relationship with some stakeholders, such as customers, suppliers, employees and competitors'* (also see Gronroos, 1996; Sheth and Parvatiyar, 1995). In addition, Hunt (1997) further argued that R-A theory can provide a theoretically grounded relational marketing, since R-A theory views intangible assets as resources and resources are heterogeneous and immobile.

Customer Relationship Management (CRM) is important for retaining existing customers. Although new market expansion is still a crucial goal in a number of retail firms, Chenet and Johansen (1999) pointed out that the cost of retaining good customers is much lower than the attraction of new customers under the same value. However, CRM may also create negative results (Davids, 1999). For example, if retailers do not consider the added value for customers, CRM will damage rather than enhance the relationship with customers.

Supplier-retailer interactions are also crucial in the retail community. A number of leading retailers may wish to maintain long-term relationships with some suppliers, since the development of the private brand product is usually time consuming. Therefore, if the contract with supplier is not stable, it will cause substantial impacts on retailer's performance. The payable turnover rate can be used to evaluate the relationship between supplier and retailer. If a retailer has good relationship with its suppliers, the time period of the payable turnover will be longer. This implies that the retailer has longer buffer time to pay the cost of the purchasing to its suppliers.

With regards to the relationship with other retailers, Dawson (2000) argued that as with increasing the size of retailer, it is more possible to have the opportunity to co-operative alliances with other competitors. In addition, small retailers may corporate together to against large retailers with the aim to survive in the market. Table 4.17 summaries some important measures of relational resources:

Table 4.17 Key Measures in Evaluating Relational Resources

Customer Relations Management	• Customer complaints management • Loyalty card strategy • Customer satisfaction • Goods returned management
Supplier Relations Management	• Good global reach • Cost sharing with suppliers on promotions • Payables turnover
Competitors Relations Management	• Co-operative alliances opportunity • The retailer association

4.3.8 External Environmental Factors

As mentioned in the Chapter Three, the R-A competition process will be affected not only by the company's internal resources, but also by external environmental factors, including societal resources on which a firm draws, societal institutions that structure economic actions, actions of competitors and suppliers, consumer behaviour, and public policy decisions. This section will discuss the main measures for evaluating such external environmental influences.

4.3.8.1 Actions from Customers, Suppliers and Competitors

Actions from customers, suppliers and competitors have great impact on a retail company's operation. For example, the changes in customer's tastes will cause a retailer to change its marketing and store operating policies. In addition, the changes of the contract content with suppliers will lead a retailer to modify its logistic and buying strategies.

Regarding actions from competitors, Hunt (2000: 142) stated that the rival actions which will damage a company's comparative advantage in resources, include: 1) buying the same resource to gain the advantaged, 2) imitating the resource from an advantaged competitor, and 3) innovating new resources. Thus, these actions from the external stakeholders should be taken into account for evaluating a retailer's performance.

4.3.8.2 Other External Environmental Factors

Apart from the actions from customers, suppliers and competitors, societal resources, societal institutions and public policy decisions will also effect a firm's competitive position. For example, an increasingly elderly population will change or affect a retail company's marketing strategy, since older people have different demand for goods and services compared with younger ones. This example shows how important it is for companies to monitor changes in various external or macro influences and to ensure that their current strategy is appropriate.

In this research, PEST analysis – the **P**olitical, **E**conomical, **S**ocio-cultural and **T**echnological analysis will be adopted to assess these external environment influences. PEST analysis is a useful tool to analyze the macro-environment and to identify external environmental indicators of a company. The definition of the four aspects of PEST is illustrated in Table 4.18

Table 4.18 Definition of Environmental Aspects

Harrison, J.S. (2003), Strategic Management of Resources and Relationships: Concepts and Cases	
Environmental Aspect	Definition
Political Environment	Influences and trends associated with government and other political forces, both at home and abroad.

Table 4.18 Definition of Environmental Aspects (Continued.)

Economical Environment	Influences and trends associated with domestic or global economics, such as economic growth rate, interest rates, the availability of credit, inflation rate, foreign exchange rate, and foreign trade balances.
Socio-cultural Environment	Influences and trends that come from groups of individuals who make up a particular geographic region.
Technological Environment	Influences and trends, related to the development of technologies both domestically and internationally.
Wheelen and Hunger (2004), Concepts in Strategic Management and Business Policy	
Environmental Aspect	Definition
Political Environment	Forces that allocate power and provide constraining and protecting laws and regulations.
Economical Environment	Forces that regulate the exchange of materials, money, energy, and information
Socio-cultural Environment	Forces that regulate the values, mores and customs of society.
Technological Environment	Forces that generate problem-solving inventions.

- **Political Environment**

It is important to analyze the political environment for barriers to retail industry development as these would clearly have a great impact on a retail company's performance. Barriers such as laws or regulations restricting channel size are common. Hasty and Reardon (1997) mentioned US government regulations aimed at ensuring no single group has dominant economical power. These regulations can be divided into three categories: antitrust laws, price competition laws and unfair trade practice laws.

In addition, land-use planning law will also affect a retail company's location selection policy (Burt and Sparks, 2003). Indeed, a retailer cannot perform well without a good location to run its business. Newman and Cullen (2002) argued that land-use planning policy has a significant impact on where and how people can shop and retailers can trade. In addition to regulation and law, government stability is also an important factor in the political environment. If a country does not have a stable political environment, business development can be easily disrupted.

- **Economical Environment**

The most common measure of overall economic health is the gross domestic product (GDP). In general, GDP expresses the total market value of all final goods and services produced in a country during a specific year. An increase in GDP implies that people are better off than before. In other words, people have more income and tend to spend more. It can be assumed that the situation will trigger higher sales and more profits for retailers. Other economic influences may have significant impacts on the operation of the retail industry. For example, an increase of interest rate will affect a retailer's expansion policy, since expansion requires large capital requirements. This research will also take into account these macro-economic influences for evaluating retail performance.

- **Socio-cultural Environment**

Changes in the socio-cultural environment will also influence a retailer's strategy, since the retail business directly interacts with customers. As the structure of population or other socio-culture factors change, a change of the retail business strategy is also necessary. For example, an increase in young women in full-time employment has a direct impact on the sales of fashion retailers (Newman and Cullen, 2002). Similarly, an increase in income leads to lifestyle and attitude changes (Burt and Sparks, 2003). People tend to increase the consumption in leisure activities or positioning products in order to represent their social class.

Different generations also have different shopping styles. For example, 'Generation X' has a particular definition of shopping. They do not like to buy traditional products, since such goods are look-alikes and out-of-date. Thus, a retailer whose target customers are from generation X would need to sell fashionable and unique items. A fall in the birth rate may have an effect on consumer spending, especially for the retailers who sell babywear products. A decline in death rate may increase the sales of old people products. Drawing on all these examples, it is argued that socio-cultural factors should be included in evaluating retail performance in this research.

- **Technological Environment**

Technology changes people's everyday life. For example, e-business provides customers with an additional and perhaps more convenient shopping channel. Customers can buy everything without going out thus saving time and money. In addition, new technology, such as the Point-of-Sales (POS) system, helps retailer to control their supply chain by collecting valuable marketing information (Newman and Cullen 2002).

New technology can also lead to the development of new equipments for retailers. For example, food retailers can distribute perishable food thanks to the development of air-conditioned or multi-temperature delivery trucks. Retailers can reduce operating costs and increase management efficiency by using this new technology. The measures can be employed to evaluate the four external environmental aspects mentioned above are listed in Table 4.19.

Table 4.19 Variables of Main Environmental Aspects in the PEST Analysis

Merrilees and Miller (1996), Retailing Management: A Best Practice Approach	
Political Environment	• Government levels, legislation, regulations • Awards, enterprise agreements • Industry associations • Professional associations • Trade unions
Economical Environment	• Situations of the global and local market economy • Regional economies • Situation of the retailing industry • Government policies • Sensitivity to biological issues, including the nature and origins of goods

Table 4.19 Variables of Main Environmental Aspects in the PEST Analysis (Continued.)

Socio-cultural Environment	• Changing demography in USA of the population • Overall and the workforce • Multiculturalism • International tourism • Changing roles for individuals		
Technological Environment	• Development of 'high tech' in retailing and related areas such as finance • Impact of technology on the workforce • Impact of technology on the retailer		
McGoldrick, P. (2002) Retail Marketing			
Political Environment	• Change of government • Tax policies • Employment law • Minimum wage • Trading hours restrictions • Planning guidelines • Monopoly legislation • Terms of trade codes • Bargain offer regulations • Environmental laws	Economical Environment	• GDP trends • Regional economics • Disposable incomes • Saving ratio • Interest rates • Exchange rates • Fuel costs • Employment levels • National competition • International competition
Socio-cultural Environment	• Environmental concerns • Consumerism • Changing work patterns • Income distribution • Holiday/leisure time • Exercise/sport participation • Food concerns • Levels of education • Ageing population • Delays in starting family	Technological Environment	• High-tech products • Food processing and presentation • Internet or interactive television • Electronic funds transfer • Electronic data interchange • Warehouse technology • Greener vehicles • Satellite tracking • International teleconferencing • Security technologies

4.4 Fieldwork Research: Interview with Practitioners

25 formal interviews with three interviewee groups were carried out. These groups were: the retail company's management (13 interviews), the bank managers in the business loan department (8 interviews) and the industrial analysts in the investment institutions (4 interviews). Interview transcriptions are presented in Appendix A. Most members of these three groups shared common viewpoints although some differences still exist among these three groups.

Section 4.4.1 will present the pilot study which provides the underpinnings of the formal interview process. This will be followed by results from the formal interviewing process as well as a discussion on the similarities or differences in opinions among company managers, bank managers, and industrial analysts.

4.4.1 Pilot Study

The participant in the pilot study carried out on 25 May, 2004, was previously a junior manager in the international department of a leading retail company in Taiwan. The interview was face-to-face with open questions, and recorded. It lasted 25 minutes. The primary purpose was to gain insights that could be built on (through researcher's self-evaluation) to ensure the quality of future formal interviews. Detailed pilot study transcriptions and reflections are presented in Appendix B. The following is a summary.

Reflecting on the literature (Kvale, 1996; Patton, 1990), the researcher found strengths and weaknesses with regards to the manner in which the interview was conducted. Table 4.20 is a list of relevant interview criteria found in the literature. A 5-point Likert scale was employed to evaluate the researcher's performance in each criterion.

Table 4.20 Interview Evaluation Table

	Self-evaluation results				
Interviewer Criteria	Very bad	Not good	Average	Good	Very good
A. Knowledgeable					V
B. Structuring			V		
C. Clear	V				
D. Gentle / Trustworthy			V		
E. Sensitive			V		
F. Open		V			
G. Steering				V	
H. Critical			V		
I. Remembering	V				
J. Interpreting			V		

Table 4.20 Interview Evaluation Table (Continued.)

K. Asking Open-ended Questions			V		
L. Avoiding Dichotomous Questions			V		
M. Using Presupposition Questions	V				
N. Asking Singular Questions				V	
O. Using Illustrative Examples		V			
P. Using Probes			V		
Q. Using Announcements	V				
R. Providing Reinforcement				V	
S. Neutrality			V		
T. Tape-recording Issues			V		

On the whole, the researcher's main strengths in this pilot interview were: 1) deep knowledge in the subject matter, due to a complete review of previous literatures and researcher's previous working experience in the retail sector, 2) being good in encouraging the interviewee to illustrate personal viewpoints; 3) being in control of the interview process and not exceeding the time limit agreed upon; and 4) mainly asking singular questions rather than confusing questions or multiple questions.

However, despite asking singular questions, the researcher's questions were not clear enough. Clarity appeared to be the most serious problem in the pilot interview. Unclear questions lead to unclear responses, and hence, weakened the validity of the interview. Following is an example of on unclear questions:

Interviewer	*OK, thank you for your answer... And the second question is... How do you feel about the factor of "location"?*
Interviewee	*Location for??*
Interviewer	*For the store...*
Interviewee	*Yes... What should I think of location in terms of what?*

As a result, researcher wasted a lot of time in explaining ambiguous questions and still cannot obtain the information needed for the research. Moreover, the researcher scored low in terms of recalling what was said earlier, using announcements and using presupposition questions. All of these elements could have made the interview questions more interconnected and allowed the whole interview to flow better. The results would have been more comprehensive, if more attention had been paid to these elements.

In terms of being open, although the researcher tried to apply the concept, it could have been better. For example, in the researcher's opinion, market share is a factor for measuring performance, but the interviewee said, '*it's a kind of consequence. It's a result more than a factor for measuring performance.*' However, researcher did not follow up on the interviewee's opinion. Instead, he tried to pull the interviewee over to his point of view. The interviewee had to repeat again what she thought, and still, researcher did not seek to follow up but jumped to the next question. Therefore, it is a need to be more open to new ideas and follow up on them.

With regards to the using of illustrative examples, the researcher attempted to use illustrative examples to clarify interview question, but did not give very good examples. A good illustrative example should include opposing extremes so that the interviewer does not lead the interviewee. For example, a leading way of using illustrative examples in the interview would be:

Interviewer	*If two companies have the same market share, but maybe one company is more risky than the other. Do you think there is something we have to consider?*

A non-leading way of using illustrative examples would be:

Interviewer	*I've heard some people say that market share is important for measuring performance, but others say that market share is not important for measuring performance. What is your opinion on this?*

In summary, before and during the interview, the researcher should consider each criterion carefully. Moreover, more preparation is needed on designing the questions, and practicing how to conduct the interviewing process. Finally, apart from the criteria in the table above, researcher discovered from this pilot interview that many questions needed some reflection from the part of the participant. They are not issues that most people would think about everyday. It would be wise to provide some general questions in advance to the interviewee so that the interviewee has time to prepare the answers.

4.4.2 Viewpoints from the Retail Company's Management

4.4.2.1 Internal Resources: Financial Resource (Financial Scale)

Apart from the traditional financial analysis, many retail managers pointed out that the financial scale based on sales or profits can ensure a retailer's market position. As a result, most retailers view the sales and profits creation as the primary annual objective. One pharmacy store manager mentioned the strategies to create sales and profits in his company.

Regarding sales creation, the most important measure is the sales per store per day (PSD). How does a retailer increase PSD? PSD can be improved by increasing two measures: customer visits and average transaction size per visit. The manager pointed out that in order to increase customer visits and average transaction size per visit, retail companies usually sell some popular products, which are accepted by most customers. They call these products *National Brand (NB)* products. Since NB products have very good brand image in the customers' mind, NB products can enhance the customer visits and increase the average transaction size. Another method to increase these two measures is changing the product mix in order to satisfy the demand of local people. For example, in Scotland, retailers usually sell Scottish goods, such as, oat cakes, since customers prefer to buy these products, which are closer to their normal life. Finally, creating other sales channels, such as e-business, is also another method to increase a retail company's sales.

With regards to the profits creation, the most important method is to sell products with higher gross margin. They name these products *Private Brand (PB)* products –

that is, brands owned by the retailer rather than the manufacturer. Many advantages are related to selling PB products. For example, there is no fair market price of PB products and pricing policy of PB product price is more flexible. Moreover, the cost of PB products is lower than other products and hence PB products are more profitable. However, the sales of PB products depend on the brand image of the retail company. If the retailer's brand image is weak, there will not be a good performance for the sales of PB products. As a result, retailers usually need to design the NB and PB product mix in order to achieve the objectives of sales creation and profits creation. The important measures related to financial scale are illustrated in Table 4.21:

Table 4.21 Key Measures in Evaluating Store Operation (Retailer Viewpoint)

- Sales per store per day (PSD)
- Customer visits
- Average transaction size per visit
- The number of national brand products
- The number of local brand products
- The number of private brand products
- Other sales channel (internet)

4.4.2.2 Internal Resources: Physical Resource (Store Expansion)

In many retail companies, the department of store development plays an important role, since a large store number can create channel advantages, such as store fixed cost reduction. Moreover, since store number is one of the key measures of market share, a retail company with a large number of stores usually has greater bargaining power vis-à-vis its customers and suppliers.

How can retail companies enhance their channel advantage? One important method is through a sound franchise system. The most important advantage of a franchise system comes from the speed of expansion. In addition, a retail company's human resource cost and rental cost can be reduced. The success of the franchise system depends on the performance of the franchiser. No company will want to be a franchisee in a poorly performing retail company.

Another method for store expansion is acquisition. There are many motivations that makes a retailer wants to buy other retailers. For example, if a retailer intends to access a new market, one possible strategy is to buy an existing company in the new market. Compared with the franchise system, the primary drawback of acquisition is that—it is costly. The choice of the retail store expansion strategy depends on the retailer's objective.

4.4.2.3 Internal Resources: Legal Resource (Brand Strength)

If a company has a well developed brand image, there will be many advantages. For example, a good brand image makes it easier to create good relationships with customers, suppliers and other stakeholders and hence, there will be comparatively less business operation hurdles. Moreover, a good brand image will increase potential profits, since it is possible to enhance customer loyalty and to inspire customers to consume. For instance, profits can be increased by selling private brand products, as private brand products usually have higher gross margin. How does a retailer to increase the sales of private products in a retail company? By having a good brand image.

A good brand image can be created in many ways, for example by ensuring product quality and good service. Recently, the issue of social responsibility has become an important factor of developing retail brand image, as it can enhance the customer loyalty. In practice, retail managers usually measure a retail company's brand strength by referring to market surveys. From these surveys, they can understand a retail company's market position and other important market information. The brand strength measures are summarized in Table 4.22:

Table 4.22 Key Measures in Evaluating Brand Strength (Retailer Viewpoint)

- The sales of private brand products
- The image of product quality
- The image of social responsibility

4.4.2.4 Internal Resources: Human Resource

Human resource quality is very important to a retail company's performance. One retailer mentioned that an employee with good performance usually has eight characteristics: (see Table 4.23)

Table 4.23 Characteristics of Good Human Resource Quality

- The ability to achieve work objectives
- The ambition to expand the job content
- Communication skills
- The relative professional job skill and knowledge
- Responsibility
- Work attitude
- Team work ability
- Interaction with customers

The most important factor is the interaction with customers, since retailers have to directly face the customers by definition. Thus, most retailers prefer to hire store staffs with a smiling face and friendly personality. However, for an e-business retailer, the most important factor is the team work ability. Unlike other retailers, e-business retailers usually prefer to hire people with experience, since they need new staffs to contribute to the business line immediately. Often, such professional people may encounter difficulty working with other members of staff. Therefore, team work ability is the most important consideration for human resource management in the e-retailer industry.

It is also necessary for an employee to have the ability and the will to achieve company's objective. If one of the criteria is lacking, a staff will not perform well. How does a retailer enhance the quality of its human resources? It depends on a good education and training system. Thus, a number of retail mangers view the retailer's employee training system as a very important performance consideration.

Another important factor relative to human resource management is employee loyalty. As high staff turnover is one of the characteristics in the retail industry, in practice, retailers adopt the average tenure of the employees to evaluate a company's employee loyalty. They argue that the higher the average tenure is, the higher the employee loyalty is. Hence, increasing the average tenure is also an important objective for retail human resource management. One manager mentioned that the most important method to increase the average tenure is to encourage their staffs to become involved in the primary business activities. The main reason is that management can understand the thoughts from the employees through these activities and assist their employees to achieve their career objective. The important human resource management measures are summarized in Table 4.24:

Table 4.24 Key Measures in Evaluating Human Resource Management (Retailer Viewpoint)

- The quality of human resource
- The training and education system
- The loyalty of employee
- The employee's average tenure

4.4.2.5 Internal Resources: Organizational Resource (Growth Power)

The most important measure to evaluate growth power is the growth rate of operating income. A company's annual net profits usually come from two different sources: operational profits (primary business line) and non-operational profits (such as, interest income, rent income or investment income). Since the operational income comes from the company's primary business line, the growth rate of the operating income is a very important performance measure. In addition, the growth rate of sales, the growth rate of gross profits and the growth rate of net profits are also crucial and a number of retailers set their annual objective by using these three measures. Obviously, growth power is very important in retail operation.

4.4.2.6 Internal Resources: Organizational Resource (Productivity)

With regards to the employee productivity, retailers measure it in terms of: sales per employee and sales per human resource cost. One e-retailer manager pointed out that the measure of sales per human resource cost is more accurate, since it measures the productivity of sales in terms of each dollar spent on the human resource.

4.4.2.7 Internal Resources: Organizational Resource (General Execution Ability)

No company can run well without sound management. How can one examine the ability of a retail company's management to run the business? It begins with examining the internal control mechanism. A good internal control mechanism can increase the operating efficiency and reduce operational mistakes. If a retail company does not have any internal regulation, it will be problematic. For example, it is difficult to deal with the following problem: how can routine work continue when a staff member is to leave the company soon? If a company has a very good internal management system, the impact will be lower. In addition, the internal control mechanism can reduce the damage from the influences of human emotion. Finally, if a retailer does not have a consistent store operation procedure, it will also damage the store service quality. The main reason is that each customer may face different services from different stores. The inconsistent quality of service will reduce the customer's loyalty and have negative impacts on a retailer's sales.

The annual objectives achievement rate is the most common measure for evaluating a retail company's management ability. Many retailers set a higher standard in order to encourage staff to achieve higher performance. Another measure is the project performance rate. It evaluates the project performance based on the projects completion progress during a specific time scale. The project performance rate can help them to control the quality of each project and ensure the project progress will not be delayed. Finally, the acquirement of the *International Organization for Standardization (ISO)* can also be considered to evaluate a retailer's management ability, since ISO indicates the quality of the retail management system. The important measures of the general execution ability are illustrated in Table 4.25.

Table 4.25 Key Measures in Evaluating Execution Ability (Retailer Viewpoint)

- Complete internal regulations
- The annual objectives achievement rate
- Project performance rate
- The acquirement of the international organization for standardization status (ISO)

4.4.2.8 Internal Resources: Organizational Resource (Organizational Management)

A retailer needs a flexible organization due to rapid changes in the consumer market. One department store manager mentioned that the learning organization has become a popular management vision that encourages individuals to follow consumer demand and make changes to accommodate it. In a learning organization, managers must have the ability to listen and delegate. This means the communication between management and employees can become more flexible and efficient. Moreover, team work is also important. One manager pointed out that *'we encourage different departments to corporate with each other, since we believe a project with different points of view is more complete'*. Table 4.26 is an arrangement of important organizational management factors:

Table 4.26 Key Measures in Evaluating Organizational Management (Retailer Viewpoint)

- Empowerment
- The listening ability of manager
- Team work ability

4.4.2.9 Internal Resources: Organizational Resource (Inventory Management)

Inventory service control is very important in the retail industry. If a retailer usually faces out-of-stock situations, a decrease of the customer's buying intention is expected. One retail manger argued that *'One of my company's policies is that we do not allow any out of stock situations. Even if there are two or three stocks on the shelf; we still think it is an out of stock situation.'* Obviously, inventory management plays a critical role in the retail operation.

The most common measure for evaluating retailer's inventory management performance is the inventory turnover. High inventory turnover usually implies the company's goods unsalable risk is low. If a retailer does not manage its inventory well, it will lead to the situation of *'shrinkage'*. Shrinkage implies the loss of sales. A UK retail manager mentioned that there are two different types of loss in a store relative to inventory management, which are *'known losses'* and *'unknown losses'*.

Known loss usually refers to waste costs. For example, if a store cannot sell out its inventory, these overstocked goods will be thrown away and become a store's loss. It can be checked and recorded every day. Known loss is very difficult to control, since it is very difficult to predict the demand of customers. However, known losses can be improved by a good customer information system, such as, the Point of Sales (POS) system.

Unknown loss is the loss that cannot be checked immediately. It can be calculated after stores count their inventory during a period, such as every six months. After a store counts its current inventory, the store can find how many products they should have sold but did not sell. In other words, these goods have disappeared. There are many possible explanations of this situation and the most common answer is that these products were stolen by staffs or customers. Unknown loss can be reduced by a good in-store security system and hence, unknown loss is controllable. In many retail companies, unknown loss is the most important measure to evaluate a store manager's performance. If a store's unknown loss is very high, this store manager is highly likely to be dismissed. Drawing on the above, inventory management is significant, since it can ensure sales will not be eroded. The main considerations of inventory measurement are expressed in Table 4.27:

Table 4.27 Key Measures in Evaluating Inventory Management Ability (Retailer Viewpoint)

- Out-of-stock control
- Inventory turnover
- Known loss
- Unknown loss

4.4.2.10 Internal Resources: Organizational Resource (Logistic Management)

A retailer usually has a large range of suppliers. If a retailer has to deal directly with all its suppliers, the operation is very complex, costly and time-consuming. The distribution centre plays a critical role to assist retailer to solve this problem. The advantages of a distribution centre include: simplifying the store operation, maintaining the store service quality and creating synergy between retailers and suppliers.

4.4.2.11 Internal Resources: Organizational Resource (Product Innovation)

Retail products usually have a common characteristic—short product life cycle, since consumer tastes change very fast. If a retailer has good ability to develop new products to quickly satisfy customers' demand, it will ensure a retailer's competitive position. Many measures can be employed to evaluate the product innovation ability of a retail company. For example, the amount of new products introduced in a specific time period can be used to measure a retail company's product innovation ability. If a retailer has more new products than other retailers within the same time period, then this retailer has better product innovation ability. Moreover, the life of new products is also important. If a new product has long life, then the popularity of the new product is sound. Furthermore, it also implies the retailer has better ability to know customer's demand and to develop new products.

Finally, the speed of new products development is also crucial. It includes two parts: the speed of new products introduction and the speed of elimination of dead items (items with low sales). A chain store retailer's CFO mentioned that two growth measures could be used to evaluate the performance of new product innovation:

- ((The number of new products introduced at T_1 – The number of new products introduced at T_0)) / The number of new products introduced at T_0) * 100%
- ((The number of eliminated products at T_1 – The number of eliminated products at T_0)) / The number of eliminated at T_0) * 100%

He further pointed out that these two ratios had better be equal to or greater than zero. If these two ratios are negative, this may imply that a retail company decreases its efforts to examine the demand of target markets. The important product innovation ability measures are expressed in Table 4.28:

Table 4.28 Key Measures in Evaluating Product Innovation Ability (Retailer Viewpoint)

- The amount of new products introduced in a time period
- The life of new products
- The speed of new products development
- The speed of elimination of dead items

4.4.2.12 Internal Resources: Organizational Resource (Marketing Management)

The most important retail marketing strategy is the differentiation strategy. Every retail company has to know what its market position is and what its strength is. Many different differentiation methods can be employed in the retail marketing management, such as products differentiation, services differentiation, or store layout differentiation. One department store manager mentioned the differentiate strategy in her company that:

> *'Our marketing strategy is to give* ***surprises*** *to our customers. Therefore, we try to differentiate our store image and layout in order to achieve the marketing objective.'*

In addition, promotion activities are also very important in the retail industry. For example, the revenue from the seasonal promotions, such as Christmas promotion, usually comprises the largest part of the annual sales in a retail company. Apart from the promotion activities, changing the store layout and remodeling are also very significant marketing activities, since they can inspire customers' buying desire. As mentioned in the Section 4.3, customer relationship management (CRM) is also an important strategy in retail marketing management. A number of retailers develop CRM by issuing customer card in order to enhance the customer loyalty. Table 4.29 arranged the important marketing performance measures:

Table 4.29 Key Measures in Evaluating Marketing Management Ability (Retailer Viewpoint)

- Differentiate strategy
- Promotion activities
- The frequency of store layout changing
- The frequency of remodelling
- Customer relationship management

4.4.2.13 Internal Resources: Organizational Resource (Technology Management)

As mentioned before, due to the rapid change of customer demand, the ability to collect customer information is imperative. Technology support systems, such as POS systems, can help retailers to collect information regarding the sales structure and the customer structure. For example, POS systems can assist retailers to answer the following questions: What are the most popular products during Christmas? Who are our target customers? What is the customer age structure of this new product?

Apart from the market information collecting system, the management support system is also important. For example, due to a huge range of products in a store, retail firms need establish a sound supply chain management system. Furthermore, as retail firms usually have numerous stores, it is time-consuming and costly to collect accounting data from each store. Therefore, a strong accounting support system can also enhance the financial operating performance. The key measures relative to retail technology management are illustrated in Table 4.30.

Table 4.30 Key Measures in Evaluating Technology Support Ability (Retailer Viewpoint)

- The investment of technology
- The strength of data collection system
- The strength of data process system

4.4.2.14 Internal Resources: Organizational Resource (Financial Management)

Most retail managers pointed out two important issues related to financial management: expense control and cash flow management. Regarding expense control, the most important expense in a retail company is wages. One UK hypermarket store manager argued that '*there are almost 500 employees working in a UK hypermarket and their wage comprises almost 6.5 percent of total sales. If a retail store cannot control the labour cost well, it will see that as having a very serious impact on its profits. Therefore, if I want to examine a store's performance, I will first examine the performance of wage management.*' Human resource cost can be reduced by increasing the percentage of part-time staffs. Therefore, retail management will evaluate the employee structure in order to measure performance.

Another important financial issue is cash flow management. If a company invests its cash flow in fixed assets, such as POS system, it will enhance a company's future operation. However, if a retail company intends to utilize its cash flow for long-term investment, this company needs to consider the synergy with the target company. For example, if a retailer invests in a distribution centre company, the investment will create synergy due to a strong corporative relationship between them. On the other hand, if a retailer invests in a mining company, there will not be any benefits to either company, since the retail company does not have any know-how to operate a mining company. Table 4.31 is an arrangement of important retail financial management measures.

Table 4.31 Key Measures in Evaluating Financial Management Ability (Retailer Viewpoint)

- Expense control
- Part-time staff ratios
- Cash flow operation strategy

4.4.2.15 Internal Resources: Informational Resource (Strategic Vision)

Every industry has its own business cycle. Different phases in the business cycle need different business strategies. As a result, understanding the phases of an industry's business cycle is also important for measuring retail performance. Most retail managers argued that strategic vision is important, since it helps the retailer face challenges in different phases of the business cycle. Moreover, strategic vision also creates a stable working environment. For instance, a clear strategic vision can enhance internal cohesion and operating efficiency. As a result, all employees will work in the same direction in order to achieve the company's objective.

4.4.2.16 Internal Resources: Relational Resource

One retail manager mentioned that *'Team'* is the key for retail performance. Team includes all stakeholders in the retail business operation, such as internal staff, suppliers and customers. A good team can create powerful synergies in the whole value chain. The performance of the team depends on the relationship among team members. As a result, the relational resource is very important in the retail business. Resources-sharing is the most common method to enhance the relational resource. For example, a retailer can cooperate with its suppliers to reduce the cost for a promotional activity. Moreover, management can share knowledge with staff with the aim of continuous learning. One chain store company CFO pointed out, *'I believe that company shares its resources with other stakeholders will have better performance.'*

4.4.2.17 External Factors: Political Environment Impact

The most serious influences from the political environment are retail industry regulations. If government limits the time period of new store development, the channel development plan will be hindered. For example, a CFO in a Philippine chain store company mentioned that if a retail company needs three months to open a new store, it is impossible to create channel advantage. He also pointed out that an efficient infrastructure system is important. For example, if the distribution company cannot transport goods on time, the quality of the inventory service level is poor.

4.4.3 Viewpoints from the Bank Managers in the Business Loan Department

4.4.3.1 Internal Resources: Financial Resource

Lenders usually assess a retail company's financial performance through several financial ratios. The most important measure is the gross margin. A Taiwanese banker pointed out two reasons to evaluate a retail company's gross margin. The first reason is to understand the bargaining power of the retail company vis-à-vis its suppliers. If a retail firm has a higher growth margin than average, then the bargaining power vis-à-vis its suppliers is also higher. The second reason is to understand a retail company's future profitability. Given that one of the characteristics of a retail industry is its low margin, having higher gross margin than other retail companies would imply higher competitiveness in the future.

Most lenders also pay attention to the financial scale. They prefer to lend money to large size firms. They believe that large companies have better stability and sustainability. Several measures can be used to evaluate retail financial scale, such as, sales, store numbers and profits. A banker pointed out that the most important measure is the sales, as most retailers sell low margin products and hence sales expansion is the only way to enhance profits in the retail industry.

4.4.3.2 Internal Resources: Physical Resource (Reach Ability)

Different retail formats need different locations, since their target customers are different. For example, department stores are usually located in a high population density area, whilst hypermarkets are usually located in a place with convenient transportation and parking. Some lenders pointed out that they will consider the retailer's location prior to the loan decision. The store number is also an important measure for a loan decision. A retailer with large number of stores usually has the advantage of economies of scale. Economies of scale can reduce the fixed costs in each store and increase the potential profits. Moreover, retailers with large number of stores usually have larger bargaining power to their suppliers than other companies.

4.4.3.3 Internal Resources: Legal Resource (Brand Strength)

Brand strength can be regarded as an intangible asset of a company. It also implies the strength of customer's loyalty. Lenders usually prefer to lend money to a company with good brand image, since they think this company has been accepted by most customers. How can lenders evaluate a retailer's brand strength? They usually measure a retailer's brand strength from some secondary materials, such as the market survey. A number of business magazines carry out market survey regarding the brand image of retail firms every year. From these surveys, lenders can understand the rank of each retailer's position in terms of the brand strength.

4.4.3.4 Internal Resources: Organizational Resource (Product Innovation)

A good retail company not only understands customer demand, but also leads customer demand. Its product innovation ability plays an important role in leading the market demand. Lenders usually prefer to lend money to a retailer with popular products. They pointed out that the sales situation in these companies will be more stable than other retailers. Thus, product innovation ability is also an important variable for lenders to evaluate a retail company's performance.

4.4.3.5 Internal Resources: Organizational Resource (Debt Repayment Ability)

With regards to the issue of the debt repayment ability, lenders usually evaluate the debtor's previous credit history first. They can obtain the credit history information from credit evaluation institutes, such as the Joint Credit Information Centre[1] (JCIC) in Taiwan. They argued that it is very difficult to lend money to a company without good credit history. However, more key issues should be considered. For example, an UK banker pointed out that *'I think market information is more important than the debtor's previous credit history, since market information is richer than other secondary information.'* If a debtor has a good credit situation in the past, but its market reputation is poor, they will reconsider the loan decision carefully. They usually obtained market information through the contacts with other bankers. They tend to join a number of banking social activities for the purpose of obtaining the market information.

Aside from the evaluation of past credit history, lenders also measure the retailer's sales growth. That a retail company's revenue is dropping implies that the ability of the company to repay the loan is also declining. Moreover, they also argued that they

---

1 JCIC Website: http://www.jcic.org.tw/index2.htm

prefer to lend money to publicly listed firms than private firms, as they believe that listed companies have more stable loan repayment ability. Other measures, such as account receivable turnover, account payable turnover and inventory turnover, are also important to evaluate a retail company's repayment ability. In other words, lenders also try to understand the pressures from the cash flow operation in a retail company before they make a loan decision. Finally, lenders usually ask for mortgage targets, repayment insurance or repayment promise from the retailers before they make a loan decision. All these factors can be regarded as a guarantee of a company's repayment ability. The important measures of a company's repayment ability are summarized in Table 4.32.

Table 4.32 Key Measures in Evaluating Loan Repayment Ability (Lenders Viewpoint)

- Debtor's past credit history
- Market information
- Sales growth situation
- Mortgage targets, repayment insurance and repayment promise
- Stockholder's background

4.4.3.6 Internal Resources: Organizational Resource (Cash Flow Management)

Retail companies usually have more cash flow than other industries, since most retailers receive cash form customers directly. Thus, the use of the cash flow is also an important consideration for lenders to make a loan decision. Lenders prefer to lend money to a company that plans to invest its cash flow to the primary business line, since these investments can assist in this company's future development. In contrast, if a retailer intends to invest its cash flow for an arbitrage purpose, lenders will view these situations as negative signals. For example, if a retailer asks for a loan with a lower interest rate with the aim of repaying a previous loan that has a high interest rate, the lender will consider the loan decision more carefully.

4.4.3.7 Internal Resources: Informational Resource (Strategic Vision)

Before lenders make a loan decision, they usually carry out interviews with retail management. Through the interviews, lenders can understand the strategic vision of a retailer. For example, a Taiwanese lender pointed out that as Taiwan local market is limited, and a retailer with good strategic vision should have future international expansion plans. If a retail company has a stable future expansion plan, they view it as a positive factor for the loan decision.

4.4.3.8 Internal Resources: Relational Resource

A number of lenders pointed out that they prefer to lend money to a retail company with a strong stockholder background, since it implies a strong financial support behind the loan. For example, Uni-President group is the largest food manufacturing group in Taiwan. If Uni-President group plans to invest in a retail firm in order to create its retailing channel, most Taiwanese lenders will possibly lend money to this new retail company.

4.4.4 Viewpoints from the Industrial Analysts in the Investment Institutions

4.4.4.1 Internal Resources: Financial Resources

Investment institutions adopt numerous quantitative measures to evaluate a retail company's performance, such as the gross margin, current ratio and operating expense rate. Apart from the accounting ratios, they also consider the market measure—particularly the Price-Earnings Ratio (P/E ratio). They argued that based on an investor's point of view, lower P/E ratio usually implies higher potential future return. However, a company with a higher P/E ratio does not always show a negative signal, since it also indicates a better growth power.

For example, Hi-tech industry usually has higher P/E ratio than traditional food manufactures. Therefore, it is necessary to examine the P/E ratio carefully. The choice of the P/E ratio depends on the purpose of investment and it varies in terms of different industries. One analyst pointed out that the appropriate P/E ratio can be estimated by referring the average P/E ratios from other foreign retail companies in order to do a more objective analysis.

4.4.4.2 Internal Resources: Physical Resources

As with retailers and lenders, investors also view store location as an important retail performance factor. The most frequent used measures are: store number, number of customers per day and average transaction size. They can obtain the information by attending the operation presentations in each retail company.

4.4.4.3 Internal Resources: Organizational Resources

The completeness of a logistics system and inventory management is also a significant factor for industrial analysts to evaluate a retail company's performance. They evaluate the performance of logistic and inventory management by assessing the inventory service level and the most important measure for this is the inventory turnover. Industrial analysts also pointed out that the ability of financial management is important. They usually use the expense to sales rate to assess the performance of expense control. Moreover, investors do not prefer to invest in a retail company with high levels of cash outflow, since this company's operating risk is high. However, if a retail company invests cash flow in its primary business line, this company will have long-term benefits. Under this situation, cash outflow may not have negative impacts on investment.

4.4.4.4 Internal Resources: Informational Resources

All the interviewees argued that they will invest in retailers with a sound long term strategic vision. An analyst argued that a good retailer usually has future international expansion plans, since this company knows that the demand of the domestic market will be saturated one day.

4.4.4.5 External Factors: Macro-Economics Factors

One analyst mentioned that before they make an investment decision, they will evaluate the external environment, especially for the macro-economic environment.

For example, if the building industry booming, then the sales of the DIY furniture retailers will be expected to increase. Thus, the external environmental factors are also important to evaluate a company's performance.

4.4.5 Discussion of Interviews

Previous sections summarized the interview results in terms of different interviewees. Following are the discussions relative to the common and different opinions among stakeholders.

4.4.5.1 Common Opinions among Stakeholders

Regardless of loan or investment decision, financial scale (in terms of sales or profits) is the most important consideration, since it presents the stability of retail operation. Moreover, the enlargement of financial scale, such as the expansion of PSD is also crucial among three groups. Besides, brand strength also plays a vital role in the retail business. Most interviewees argued that a retailer with good brand image is easier to create good relationship with outside stakeholders. Therefore, brand strength has the advantage of reinforcing the efficiency of operation.

With regards to the physical resource, all three interviewee groups pointed out that the amount of stores is imperative, since a large number of stores can create channel advantages, such as fixed cost reduction and enhancing the bargaining power with suppliers. On the subject of the organizational resources, the common opinions among three groups are: product innovation ability, inventory management, logistic management, technology development and financial management. For example, regarding the financial management, all three groups pointed out that the cash flow management and the expense control are the most important considerations. In particular, almost all the stakeholders argued that a retail company's cash flow is appropriate to invest in its primary business line, since the investment will create long-term benefits.

Strategic vision also shows its importance among stakeholders. Retail management should have the ability to lead its company to the right direction in order to maintain or to augment its market competitive strength. Thus, understanding the industry trend and developing a high quality future development plan are also important performance considerations. Finally, all three interviewee groups mentioned the significance of the relational resource. For example, a retail manager mentioned that the performance of a retail company depends on the performance of the *'team'*, which includes all stakeholders in the retail operation. Moreover, lenders usually prefer to lend money to a retail company with strong stockholder background, as it presents a strong financial support behind the loan.

4.4.5.2 Different Opinions among Stakeholders

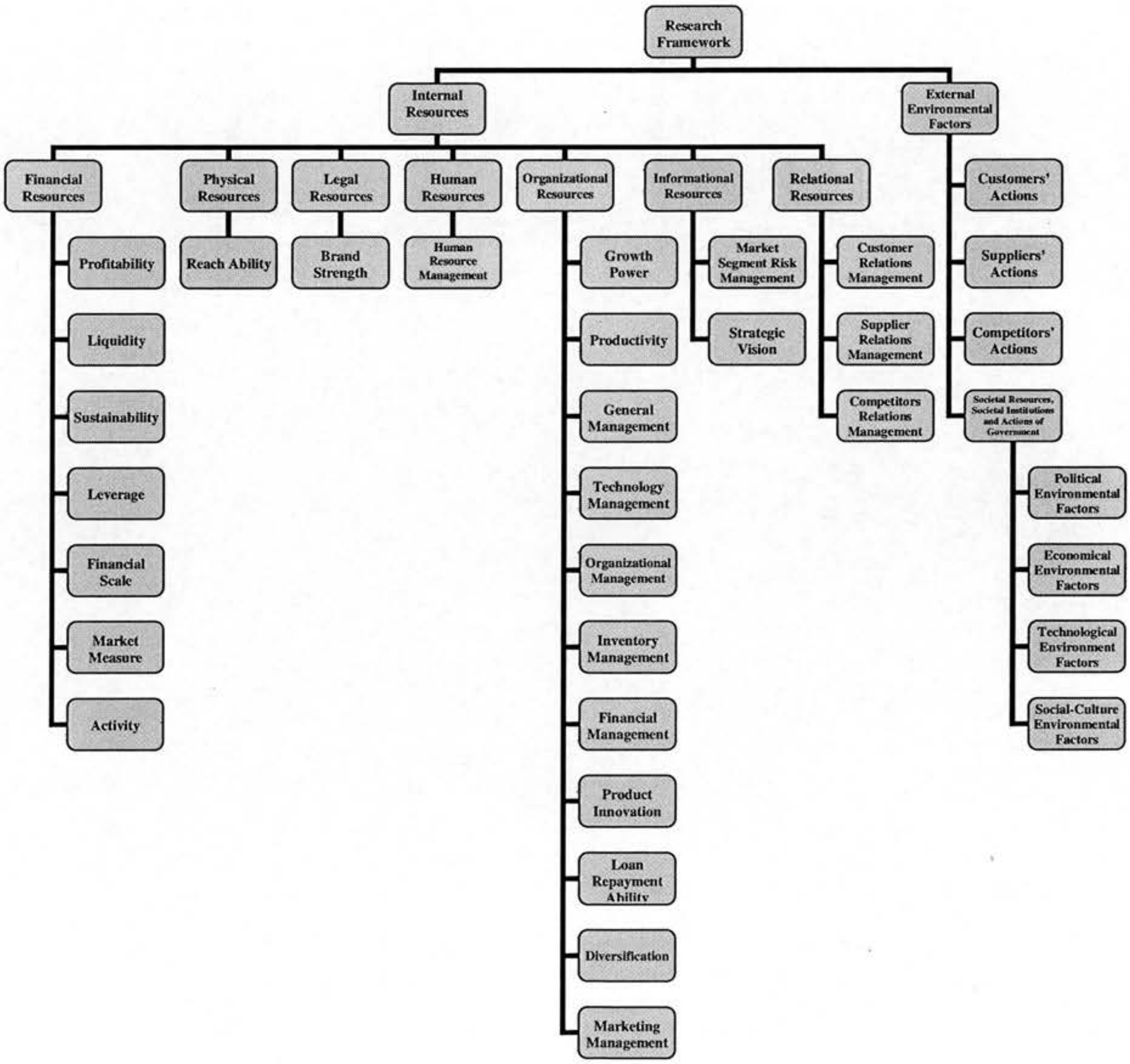
The most obvious difference amongst stakeholders is their roles of retail performance measurement. For example, for retailers, important performance measures are related to the execution ability, since these performance factors are close to their routine job. For the managers in the bank's business loan department, the most important consideration is the loans repayment ability. The only question they are concern with is: *Can we collect our loan in the future?* As a result, they usually evaluate a debtor's past credit situation and the use of the loan. In addition, they are also interested in the retail cash operating pressure, since the sufficiency of cash flow is the guarantee for a retailer to repay their obligations.

Finally, with regards to the industrial analysts, most of them mentioned that before they make an investment decision, they usually evaluate a company's performance by some market measures, such as P/E ratio. The major concern is that estimating a company's market value is key to an investment decision and P/E ratio is one of the market value estimation techniques. Drawing on above, it is obvious that different stakeholders have different opinions regarding the retail performance measurement in terms of their specific roles.

4.5 Research Framework Construction

So far, 170 retail performance measures in total were gathered both previous relevant literatures and viewpoints from practitioners. Based on the R-A theory, all performance measures can be classified into two categories: internal resources and external factors. The arrangement of these performance measures is presented in Appendix C. Furthermore, the research framework can be constructed in terms of the R-A theory and it is presented in Figure 4.2:

Figure 4.2 Research Framework



4.6 Concluding Remarks

This chapter has constructed a framework for developing a retail performance measurement system based on Hunt's (2000) R-A theory. The collection of performance measures is based on secondary materials from existing literature and primary materials from practitioners' viewpoints. The secondary materials review of performance measures is built on academic literature, reports from credit rating companies and retail management textbooks.

On the subject of interview-based fieldwork, the study carried out 25 interviews in terms of three stakeholder groups: the retail companies' management (13 interviews), the bank managers in the business loan department (8 interviews) and the industrial analysts in the investment institutions (4 interviews). It does appear that most of members of these three groups shared a common viewpoint. However, some differences still exist among these three groups. The most obvious difference among stakeholders is their roles of retail performance measurement.

Overall, 170 variables were found. According to the R-A theory, a company's internal resources can be divided into seven different types of resources: financial resources, physical resources, legal resources, human resources, organizational resources, informational resources and relational resources. All the internal performance measures were distributed to these seven types of resources. For example, the brand image of a retail company can be regarded as a legal resource, since it is protected by law and competitors cannot take advantages of it.

Regarding external factors, the actions of customers, suppliers and competitors have great impacts on retail performance. For example, the changes of customer's tastes will cause a retailer to change its marketing and store operating policies. In addition, the changes of the contract with suppliers will lead a retailer to modify its logistic and buying strategies. Apart from the actions from stakeholders, other external factors, such as the public policy decisions, also will affect a retailer's performance. A general environment analysis tool: PEST analysis–the Political, Economical, Socio-culture and Technological environment analysis was employed in order to assess the influences from the external environment. Wheelen and Hunger (2004) mentioned that PEST analysis is a useful tool to identify a company's external environmental influences in order to avoid strategic surprise and ensure its long-term health.

Drawing on above, it is obvious that this research framework has considered as many potential retail performance measures as academics and practitioners have thought of and used. However, although this research conducted 25 interviews with the aim of obtaining the viewpoints from practitioners, the importance of these variables is still unclear. Consequently, the researcher also carried out a survey in order to obtain more insights from the context. Through the survey, the researcher can quantify each variable with the intention of comparing the degree of importance among variables. It is possible that some current variables may be insignificant after conducting the survey and hence, they should not be considered in this research framework.

The primary objective of the survey can be viewed as a robust examination of the model framework. In addition, a series of comparative analysis are achievable in terms of the various insights from different countries, different retail management functions and different retail formats. In the next chapter, a discussion of all the relevant information regarding the survey analysis will be presented.

Chapter FIVE

Survey Examination of the Research Framework

5.1 Introduction

Chapter Four presented an initial research framework constructed from 170 retail performance measures found in both existing literature and mentioned during practitioner interviews. However, the importance of these variables, especially the qualitative ones, is still indistinct. It is possible that some variables may not hold great importance, despite being used before or mentioned by practitioners. To find out the level of importance of each of the variables, survey was employed. Unlike the interviews, which are restricted by resource, surveying has the advantage of studying a wider sample. In addition, surveying will not be limited by geographical restrictions. A survey will enhance the information from the context and will allow further examination of the framework.

This chapter begins by introducing the key issues related to survey design, including: sampling strategy, variable selection, questionnaire format and implementation. It moves on to insights gained from the pilot survey, carried out to improve the validity and reliability of the final version of the questionnaire. Following this, the results from the formal survey are presented through descriptive analysis and comparisons among different countries, management functions and retail formats. Next is a discussion on a case study carried out in order to explore the difference between the 'expected value' and 'actual performance' of each variable. The results facilitate the identification of the case company's competitive position based on the R-A competition theory. The final section summarized all the findings in the chapter.

5.2 Survey Design

5.2.1 Sampling Strategy

The population of the survey is defined as members of the management teams in those publicly listed retail companies. With the aim of comparing views from different geographical areas, the primary sampling strategy was *Two-Stage Geographical Cluster Sampling*. In the first stage, three target countries—UK, US and Taiwan— were selected in order to explore different opinions from Europe, North America and East Asia. The second stage was to select publicly listed retail companies in these three countries as the sample population.

Denscombe (1998) defined cluster sampling as *'Studying a small concentration of people thought to be representative of the population.'* Bryman (2001) pointed out that cluster sampling is appropriate, when the population covers a large region. For example, if the research sample is defined as 1000 students in the UK, a great deal of travel would be expected for researcher to collect data from various universities. Therefore, it is more convenient to choose a few universities first, and then select 1000 research samples from these selected universities. Obviously, the primary advantage of cluster sampling is the cost and time saving. However, Wu (2005) argued that since cluster sampling is always a multi-stage sampling approach, the possibility of sampling errors is higher than in other probability sampling methods.

5.2.2 Survey Variable Selection

As mentioned earlier, the research framework so far has taken into account 170 variables. These variables are presented in Appendix C. From Appendix C, it is clear that some measures can be easily quantified by calculating ratios or collecting accounting numbers from annual financial statements. The importance of these quantifiable variables can be evaluated by comparing their values. However, other measures, which are either difficult to quantify or are quantifiable but with no relevant data available, require further study in order to examine their importance.

Surveying is one of the techniques that can be employed to quantify these *'unquantifiable'* variables. Prior to surveying, it is useful to classify the totality of performance measures into the following groups: *'quantifiable measure and available data'* group, *'quantifiable measure but no available data'* group, and *'difficult to quantify'* group. The complete classification is illustrated in Appendix D.

Examples in the *'quantifiable measure and available data'* group would be financial ratios, such as current ratio, inventory turnover or debt ratio. These financial ratios are available in each company's financial statement and are easy to calculate. Examples in the *'quantifiable measure but no available data'* group would be part-time staff ratios or the frequency of changes in marketing strategy. Although these ratios are easy to calculate per se, they do not have to be published by retailers. In other words, researchers need to contact the company for the information. At the same time, such information can be considered as confidential by the organization and may not be easily released, even to an independent or neutral researcher. Finally, an example of the *'difficult to quantify'* group would be the quality of a company's future strategies. This is highly subjective and therefore harder to define. As a result, the variables in the *'quantifiable measure but no available data'* group, and *'difficult to quantify'* group will be the candidate variables for questionnaire design.

However, one important consideration for questionnaire design is the length control (Bryman, 2001). There are 113 variables in both the *'quantifiable measure but no available data'* group, and the *'difficult to quantify'* group. It is unwise to consider all of these variables in a questionnaire. Therefore, this research only selected 44 variables for the final questionnaire design based on R-A theory.

In R-A theory, the *internal resources* can be divided into seven categories: financial resource, physical resource, legal resource, organizational resource, human resource, informational resource and relational resource. Each type of resource has its own main principals, such as the *'reach ability'* for physical resource and the *'brand strength'* for legal resource. From Appendix D, it is obvious that almost all the principals in the financial resources group and the principal of growth power are in the *'quantifiable measure and available data'* group. These variables need not be considered in the survey. With regards to other internal resources, at least two variables were selected for evaluation of each principal and in total, 35 variables were chosen for the survey. These variables are presented in Table 5.1:

Table 5.1 Survey Variable Selection (Internal Resources)

Physical Resource	
Principal	Measures (Variable Code)
Reach Ability	1. The footfalls of major outlets (V1) 2. Trading area and store locations (V2)
Legal Resources	
Principal	Measures (Variable Code)
Brand Strength	3. The sales of private brand products (V3) 4. The image of social responsibility (V4)
Human Resources	
Principal	Measures (Variable Code)
Human Resource Management	5. Turnover (V5) 6. Staff orientation and training (V6)
Organizational Resources	
Principal	Measures (Variable Code)
Expansion Ability	7. The completeness of the franchise system (V7) 8. Store opening program (V8)
Productivity	9. Average weekly sales per square meter (V9) 10. Spend-per-visit rate (V10)
General Management	11. Internal regulations (V11) 12. The annual objectives achievement rate (V12)
Technology Management	13. The investment of technology (V13) 14. The strength of data collection and process system (V14)
Organizational Management	15. Empowerment (V15) 16. The listening ability of management (V16)
Inventory Management	17. Loss control (V17) 18. Out of stock situation (V18)
Marketing Management	19. Differentiation strategy (V19) 20. The frequency of remodelling (V20)
Financial Management	21. Cost control ability (V21) 22. Part-time staffs ratio (V22)
Product Innovation Ability	23. The life of new products (V23) 24. The speed of new products development (V24)
Loan Repay Ability	25. Debtor's past credit history (V25) 26. Stockholder's background (V26)
Diversification	27. Capital expenditures in internet channel (V27) 28. Maintaining target customer group in market diversification (V28)

Table 5.1 Survey Variable Selection (Internal Resources) (Continued.)

Informational Resources	
Principal	Measures (Variable Code)
Market Segment Risk Management	29. Following fashion trends (V29) 30. Facing seasonal demands (V30)
Strategic Vision	31. Openness to criticism (V31) 32. Willingness to innovate or experiment (V32)
Relational Resources	
Principal	Measures (Variable Code)
Customer, Supplier, Competitor Relations Management	33. Customer complaints (V33) 34. Cost sharing with suppliers on promotions (V34) 35. Joint venture opportunity (V35)

With regards to the *external factors*, R-A theory uses five groups: actions of consumers, actions of supplier-competitors, societal resources, influences from the societal institutions and actions of government. Apart from the variables within the groups focused on actions from consumers, suppliers and competitors, this research also adapted PEST analysis to evaluate the societal resources, the influences from the societal institutions and the actions of government. As most macro-economics variables such as GDP, interest rate, or unemployment rate, are quantifiable and are available from the public resources, the influences from the macro-economics environment will not be discussed in the survey. As with the internal resources, each external factor has its own principals and at least two measures were selected for evaluating each principal. Overall, nine variables were chosen for questionnaire design and they are presented in Table 5.2:

Table 5.2 Survey Variable Selection (External Factors)

The Actions of Customers, Suppliers and Competitors	
Principal	Measures (Variable Code)
The Actions of Customers, Suppliers and Competitors	36. Changes in customer's preferences or tastes (V36) 37. Changes in supplier's contract content (V37) 38. The innovation and imitation from competitors (V38)
The Societal Resources, the Societal Institutions and the Actions of Government	
Principal	Measures (Variable Code)
Political Environmental Factors	39. Change in government laws (V39) 40. Stability of government (V40)

Table 5.2 Survey Variable Selection (External Factors) (Continued.)

The Societal Resources, the Societal Institutions and the Actions of Government	
Principal	Measures (Variable Code)
Technological Environmental Factors	41. Innovation of new technology equipment (V41) 42. New management system software development (V42)
Socio-culture Environmental Factors	43. Change of population structure (V43) 44. Change of lifestyle (V44)

5.2.3 Questionnaire Format

There are two main sections in the questionnaire. The first section is called: *'Evaluating the Importance of Various Factors to the Performance of Your Retail Company'*. This section asks respondents what the key factors influencing the performance of their companies are as well as how important they are. In other words, the first section tries to measure the *expectations* from respondents regarding the importance of each key performance measure. Four questions are designed by using the five point *Likert Scale* measurements. Respondents need to pick a rating from *'Don't Know'* to *'Absolutely Important'* in terms of the degree of perceived importance for each factor. Figure 5.1 is an example of Likert scale question in the first section.

Figure 5.1 Example of Likert Scale Question in the Section One of Survey

Question 1:

The following factors are related to Reach Ability, Brand Strength, Human Resource Management, Expansion Ability and Productivity.

How important are these ten factors to your company's performance?

Please tick one box only from 'Don't Know' to 'Absolutely Important' for each statement.

	Absolutely Important	**Very Important**	**Moderately Important**	**Not Very Important**	**Don't Know**
Number of customer visits	☐	☐	☐	☐	☐
Store location	☐	☐	☐	☐	☐
Sales of the private brand products	☐	☐	☐	☐	☐

The second section of the survey is called *'Evaluating Performance through a Ranking Process'*. This section requires respondents to evaluate their companies' *'Actual'* performance in terms of the variables mentioned in the first section. For instance, the typical questions will be: how well they think their companies are doing in the factors mentioned in the first section or how much impact such factors have on their companies' performance. The five point Likert scale was also employed in this section. Respondents need to choose a raking from *'Don't Know'* to *'Extremely Good'* or from *'Don't Know'* to *'Extremely Strong'*. Figure 5.2 is an example of question in the section 2.

Figure 5.2 Example of Likert Scale Question in the Section Two of Survey

Question 8:

How strong impact the following factors have on your company's performance?

Please tick one box only from 'Don't Know' to 'Extremely Strong' for each statement.

	Extremely Strong	Very Strong	Moderately Strong	Not Very Strong	Don't Know
Changes in customer's preferences	☐	☐	☐	☐	☐
Changes in supplier's contract content	☐	☐	☐	☐	☐
The innovation and imitation from competitors	☐	☐	☐	☐	☐

In addition to Likert scale questions, open question served to obtain additional comments from respondents. Finally, in order to conduct the comparison analysis in terms of different departments, categorical questions regarding the job content of each respondent is also included in this survey. The final version of questionnaire is presented in Appendix E. As Taiwan is one of the target countries, the original English version questionnaire was also translated into Mandarin. The Mandarin version of questionnaire is presented in Appendix F.

5.2.4 Implementation

This research selected e-questionnaires as the main survey instrument. There are many advantages related to an e-questionnaire. Compared with other types of questionnaires, such as postal questionnaires, the cost of e-questionnaires is much lower. In addition, e-questionnaires are also convenient for respondents: when respondents finish completing an e-questionnaire, all they have to do is to click a button to submit. They do not need to make an additional effort to post the questionnaire back to the questionnaire administrator. Furthermore, respondents can access the e-questionnaire electronically whenever and wherever they want to. An e-questionnaire would blend in with the way people work in the digital world today.

The main argument against using e-questionnaires is that when respondents receive the e-questionnaire mail, they usually delete it. This is highly possible in very busy working environments. The results will have a negative impact on the response rate. In addition, e-questionnaires also have the shortcoming in that the respondents' feedback may be incomplete or invalid, since the researcher can never be sure whether the right person has answered the questionnaire or not.

With the aim of overcoming these drawbacks, this research sent out the e-questionnaire to the sample companies, but also followed up with each sample company by phone to request each to complete the e-questionnaire. The staff in the investor relations department or the public relations department of each sample company will be the main contacts for this research, since these two departments are usually responsible for communicating with outside stakeholders. Thus, an email was first sent out to these contact people with an attached e-questionnaire in html format as well as a link to the same e-questionnaire on-line.

A password was contained in the same email. Before respondents complete the survey, they were required to enter this password in order to validate the questionnaire. The main reason is for the researcher to identify the business code (such as, SIC code) of each sample company through this password with the intention of carrying out a comparative analysis in terms of different retail formats. After sending the email, the researcher also contacted the staff in the investor relations department or the public relations department by phone in order to obtain their agreements to forward the email to the potential respondents in their companies. Results indicated that this approach improved the response rate.

Thus far, the issues related to the survey design have been introduced, including sampling strategy, variable selection, questionnaire format and survey conducting approach. Prior to implementation, a pilot survey study was carried out in order to ensure the quality of final version of questionnaire. The next section will introduce the results and reflections on the pilot survey.

5.3 Survey Pilot Study

This pilot survey was carried out on 15, January, 2005. Twenty e-questionnaires were sent to a US retail company (Convenience Chain Store); 10 e-questionnaires were sent to a Taiwan e-business company (Internet Book Shop); 60 e-questionnaires were sent to a Taiwan retail company (Convenience Chain Store) and one e-questionnaire was sent to a UK retail company (Hypermarket).

By 20 February 2005, there were 14 responses from the US retail company (response rate: 70%), 5 responses from the Taiwan e-business company (response rate: 50%), 30 responses from the Taiwan retail company (response rate: 50%) and one reply from the UK retail company. The overall response rate was approximately 55%. Following are the main discussions and reflections about this pilot study.

5.3.1 Response Rate

The average response rate of this pilot study was above 50%. This may appear to be high, but in actual fact, most respondents were contacted using the researcher's previous personal contacts. Thus, the average response rate of this pilot study is above 50% as a matter of course. Moreover, as mentioned in the previous section, it is clear that despite personal contacts, respondents tend to delete emails containing e-questionnaires. Therefore, the improvement of the response rate was still a key issue for the survey.

Bryman (2001) mentioned several important methods for increasing the response rate. A good covering letter is very important. It can explain many significant issues relative to the survey, such as, research purposes, importance of this research, the recipient's selection criteria and the guarantee of confidentiality. In addition, a follow up procedure is also vital to increase the response rate. For those individuals who did not initially reply, two or three further emails or calls are necessary. Finally, a shorter questionnaire design can also improve response rate. One of the respondents mentioned that: *'Perhaps reassure the reader that the questionnaire is short, easy, and won't take too much time to complete right at the beginning.'* Thus, designing a short, easy-to-read questionnaire and begin with questions are that interesting to respondents will be the main objectives in the design of the final questionnaire.

A questionnaire with clear instructions and an attractive layout also improves response rate. With regards to this point, most respondents were satisfied with the design of the pilot survey, especially the instructions. Finally, providing monetary incentives will also have positive impact on the response rate. However, this research did not provide monetary incentives to the respondents, since this research would contribute feedback, such as, a retail performance report to the sample companies. Therefore, highlighting the benefits of the research to the respondents was the main strategy to encourage respondents to complete the formal survey.

5.3.2 Validity

This pilot survey adopted 44 variables to be the main performance measures. However, one may question: are these 44 variables good measures? Gilbert (2001) pointed out that validity is about whether the right concept is measured. It means whether a measure of concept really evaluates the concept (Bryman, 2001). For example, *'store locations'* and *'customer visits'* can be used to measure a retail company's *'reach ability'*. Therefore, the key point of the validity is to explore whether the concept of reach ability can be assessed by store locations or number of customer visits.

Several respondents mentioned that some terms needed to be defined more clearly, such as, *'store remodelling'*, *'operating procedures and regulations'*, *'following customer trends'* and *'high annual objectives achievement rate'*. If these terms are not well defined, they would be very difficult to measure. Therefore, this pilot study also refined these terms. Some examples are illustrated in Table 5.3.

Table 5.3 Redefinition of Terms

Old Variables	New Variables
Store remodelling	Store renovation/redecoration
Attractive store opening program	Stores opening strategy
Differentiation strategy	Market positioning
Out-of-stock situation	Inventory service level
Operating procedures and regulations	Internal procedures
Following customer's trends	Following fashion trends
Annual objectives achievement rate	Achievement of year-end goals

5.3.3 Reliability

Reliability is about whether a measure works in a consistent way (Gilbert. 2001). Bryman (2001) argued that reliability has two meanings. The first meaning is stability, meaning that if a survey is carried out on the same target group twice in different time periods, there will be little variation in the results obtained. *Test-retest* method can be used to evaluate stability by calculating the correlation between the two survey results. If the correlation is high, then the survey is reliable. However, there are some drawbacks in relation to the test-retest method. Chou (2002) argued that it is very difficult to maintain the same target group's interest to repeat the same survey in terms of different time periods. Another obvious drawback is the additional time needed to collect survey data.

The second meaning of reliability is internal reliability: whether or not the variables that make up the scale are consistent. Internal reliability can be examined by the *'split-half'* method. For example, if researchers intend to use 10 questions to measure a concept, they can divide these 10 questions into two groups via a random method. The next step is to obtain the score for each group. If the correlation between these two groups' scores is high, then so is internal reliability. The advantage of the split-half method is that researchers only collect the survey data once. Nevertheless, there are still some shortcomings about the split-half method. Chou (2002) mentioned that the split-half method requires more questions which make it difficult to design a short questionnaire. Moreover, the contents of these questions should be very similar which may reduce the respondents' motivation to complete a survey.

Nowadays, the most common method to evaluate reliability is by using Cronbach's alpha (Cronbach, 1951). Bryman (2001) mentioned that Cronbach's alpha is approximately the average of the possible split-half correlations. The formula of Cronbach's alpha is as follow:

$$\alpha = \frac{K}{K-1}\left[1 - \frac{\sum_{k=1}^{k} Var(X_k)}{Var(\sum_{k=1}^{k} X_k)}\right] \tag{5.1}$$

where K is the number of questions or variables, X is the variable value for each case and $Var(X_k)$ is the variance of the k^{th} variable.

The Cronbach's alpha will vary between 1 (meaning perfect internal reliability) and 0 (meaning no internal reliability). In general, if Cronbach's alpha is greater than 0.7, the reliability of a survey is acceptable (Chou, 2002). In this pilot study, Cronbach's alpha was 0.832. Therefore, it can be concluded that the reliability of this pilot survey was satisfactory.

5.4 Survey Analysis

Regarding the final survey, 159 e-questionnaires were sent to US sample companies and 65 e-questionnaires were sent to UK sample companies between 26 March, 2005 and 3 April, 2005. 211 e-questionnaires were sent out to Taiwan sample companies between 5 September and 12 September, 2005. The e-questionnaire could not be sent to all sample companies, since some sample companies either did not have a company website or had no apparent email address. This would be partially explained by the disappearance of these companies following bankruptcy or hostile takeovers.

After two weeks, there were only two responses from the US and 10 responses from Taiwan. In order to increase the response rate, follow up through phone calls with emails was carried out from 18 April, 2005 to 31 May, 2005 for the US and UK as well as from 26 September to 30 October, 2005 for Taiwan. Finally, there were 21 responses from US, 10 responses from UK and 120 responses from Taiwan. The response rates for US, UK and Taiwan are 13.21%, 15.38% and 56.87% respectively. The overall response rate is 34.71%. As the Cronbach's alpha was 0.864, the reliability of the final survey was acceptable.

The response situation from Taiwan clearly showed the best performance among three countries. This is mainly due to strong support from a Taiwanese leading retail company (84 responses were received from this company). With such a high response rate, it is possible to carry out a detailed case study. This case study will be discussed in Section 5.5.

There were 42 negative responses from US, 31 negative responses from UK and 15 negative responses from Taiwan. The main reason given for rejection was that based on the company policy, they do not participate in any survey. For other e-questionnaires that were sent out, this research did not receive any reply from the respondents. In next section, the focal point will focus on the first part of questionnaire in order to investigate the 'expected importance' among variables.

5.4.1 Descriptive Analysis

Descriptive analysis concentrates on two main tasks: data frequency analysis as well as the central tendency (*mean* and *median*) analysis. The original frequencies of the total 151 responses by country, department and retail format are presented in Table 5.4, 5.5 and 5.6, respectively:

Table 5.4 Original Frequency (Country)

Country Code	Country	Frequency	Percent
C1	Taiwan	120	79.47%
C2	US	21	13.91%
C3	UK	10	6.62%
	Total	151	100%

Table 5.5 Original Frequency (Department)

Department	Frequency	Percent
Accounting and Finance	34	22.52%
Investor Relations (including Public Relations)	26	17.22%
Operations	22	14.57%
Marketing (including advertising)	20	13.25%
Research and Development (Planning)	15	9.93%
Human resources	11	7.28%
Purchasing	7	4.64%
Store development (including construction franchising)	5	3.31%
Auditing	5	3.31%
Information systems	4	2.65%
Logistics	1	0.66%
Law	1	0.66%
Total	151	100%

Table 5.6 Original Frequency (Retail Format)

Retail Format (Based on the NAICS Code)	Frequency	Percent
445. Food and Beverage Stores	118	78.15%
454. Nonstore Retailers	7	4.64%
448. Clothing and Clothing Accessories Stores	6	3.97%
446. Health and Personal Care Stores	4	2.65%
453. Miscellaneous Store Retailers	4	2.65%
441. Motor Vehicle and Parts Dealers	3	1.99%
451. Sporting Goods, Hobby, Book and Music Stores	3	1.99%
442. Furniture and Home Furnishings Stores	2	1.32%
443. Electronics and Appliance Stores	2	1.32%
444. Building Material and Garden Equipment and Supplies Dealers	1	0.66%
452. General Merchandise Stores	1	0.66%
Total	151	100.00%

Regarding the frequency counts by department and retail format, it seems that the frequency for some specific areas is very small. As a result, regrouping is necessary for further statistical analysis. For the frequency count by retail format, the largest group by far is *Food and Beverage Stores*. The frequency of other retail formats does not have a comparable size. Thus, the frequency of retail format will be simply regrouped into two categories: *Food and Beverage Stores Format* and *Other Retail Formats*.

With regards to the frequency count by department, all the items are regrouped in terms of the similarity of the job content. The final largest frequencies group is the *Accounting, Finance and Auditing Departments* and *Other Departments* (including purchasing, logistics and law departments) is the lowest frequencies group. The new regrouped frequency for retail format and department are presented in Table 5.7 and 5.8, respectively.

Table 5.7 Regrouped Frequency (Retail Format)

Retail Format Code	Retail Format	Frequency	Percent
F1	Food and Beverage Stores	118	78.15%
F2	Other Retail Formats	33	21.85%
	Total	151	100%

Table 5.8 Regrouped Frequency (Department)

Department Code	Department	Frequency	Percent
D1	Accounting, Finance and Auditing	39	25.83%
D2	Operations and Store Development	27	17.88%
D3	Investor Relations (including Public Relations)	26	17.22%
D4	Marketing (including Advertising)	20	13.25%
D5	R&D and Information System	19	12.58%
D6	Human Resources	11	7.28%
D7	Other Departments (Purchasing, Logistics and Law)	9	5.96%
	Total	151	100%

In order to explore the importance of the 44 performance measures, the mean and the median were calculated based on the original Likert scale data. For example, a Likert scale with five points, such as *'Absolutely Important'*, *'Very Important'*, *'Moderately Important'*, *'Not Very Important'* and *'Don't Know'*, were coded 1, 2, 3, 4 and 5, respectively. Although mean is not an appropriate measure for ordinal scale, it was calculated for reason of assessing skewness with median.

'Not Very Important' to *'Absolutely Important'* present the different levels of importance clearly. However, the choice of *'Don't Know'* does not imply any level of importance, since there are many possible reasons behind this answer. If researcher intends to calculate an average importance of a specific variable, the average value may possible be affected by the value of the choice of *'Don't Know'*. As a result, all the original data of *'5'* will be defined as the missing value and will not be considered in calculating the mean or median. The overall ascending ranks of the mean and median values are presented in Tables 5.9 and 5.10:

Table 5.9 Rank of Mean Values (Overall)

Variable Code	Performance Measure	Mean
V2	Store location	1.42
V32	Willingness to innovate	1.44
V1	Number of customer visits	1.47
V33	Customer complaints management	1.50
V19	Market positioning	1.52
V30	Facing seasonal demands	1.61
V7	Franchise system	1.67
V9	Sales per store	1.68
V18	Inventory service level	1.68
V6	Staff training	1.70
V10	Spending-per-visit rate	1.72
V44	Change of lifestyle	1.75
V8	Store opening strategy	1.79
V36	Changes in customer's preferences	1.85
V14	Quality of data collection and process system	1.86
V29	Following fashion trends	1.86
V21	Expense control ability	1.91
V11	Internal procedures	1.94
V12	Achievement of year-end goals	1.95
V24	Speed of new products development	1.98
V28	Maintaining target customer group in market diversification	1.98
V31	Openness to criticism	2.00
V16	Response to staff issues	2.03
V41	Change of population structure	2.03
V40	Innovation of new technology equipment	2.09
V15	Empowerment of staff	2.12
V39	Change of government laws	2.13
V17	Inventory loss control	2.18
V38	The innovation and imitation from competitors	2.21
V23	Shelf-life of new products	2.24
V34	Cost sharing with suppliers for promotions	2.26

Table 5.9 Rank of Mean Values (Overall) (Continued.)

Variable Code	Performance Measure	Mean
V20	Store renovation/redecoration	2.30
V42	Stability of government	2.34
V4	Social responsibility	2.36
V25	Past credit history	2.36
V26	Financial support from stockholders	2.36
V13	Investments in technology development	2.38
V3	Sales of the private brand products	2.40
V5	Employee turnover rate	2.40
V37	Changes in supplier's contract content	2.43
V43	New management system software development	2.44
V27	Internet channel development	2.45
V35	Joint venture opportunity with competitors	2.72
V22	Percentage of part-time staff	2.92

Table 5.10 Rank of Median Values (Overall)

Variable Code	Performance Measure	Median
V1	Number of customer visits	1
V2	Store location	1
V19	Market positioning	1
V32	Willingness to innovate	1
V33	Customer complaints management	1
V3	Sales of the private brand products	2
V4	Social responsibility	2
V5	Employee turnover rate	2
V6	Staff training	2
V7	Franchise system	2
V8	Store opening strategy	2
V9	Sales per store	2
V10	Spending-per-visit rate	2
V11	Internal procedures	2
V12	Achievement of year-end goals	2
V13	Investments in technology development	2
V14	Quality of data collection and process system	2
V15	Empowerment of staff	2
V16	Response to staff issues	2
V17	Inventory loss control	2
V18	Inventory service level	2
V20	Store renovation/redecoration	2
V21	Expense control ability	2
V23	Shelf-life of new products	2
V24	Speed of new products development	2
V25	Past credit history	2
V26	Financial support from stockholders	2
V27	Internet channel development	2
V28	Maintaining target customer group in market diversification	2
V29	Following fashion trends	2
V30	Facing seasonal demands	2
V31	Openness to criticism	2
V34	Cost sharing with suppliers for promotions	2
V36	Changes in customer's preferences	2
V37	Changes in supplier's contract content	2
V38	The innovation and imitation from competitors	2

Table 5.10 Rank of Median Values (Overall) (Continued.)

Variable Code	Performance Measure	Median
V39	Change of government laws	2
V40	Innovation of new technology equipment	2
V41	Change of population structure	2
V42	Stability of government	2
V43	New management system software development	2
V44	Change of lifestyle	2
V22	Percentage of part-time staff	3
V35	Joint venture opportunity with competitors	3

From Tables 5.9 and 5.10, it is obvious that the top five variables are: *Store Location (V2), Willingness to Innovate (V32), Number of Customer Visits (V1), Customer Complaints Management (V33)* and *Market Positioning (V19)*, as well as the bottom five variables are: *Changes in Supplier's Contract Content (V37), New Management System Software Development (V43), Internet Channel Development (V27), Joint Venture Opportunity with Competitors (V35)* and *Part-Time Staff Ratio (V22)*.

As the coding of *'3'* means *'Moderately Important'* and the coding of *'4'* means *'Not Very Important'*, it can be argued that if the value of mean or median is greater than three, then the performance variable is not critical to the retail operation. Since Tables 5.9 and 5.10 show no value greater than three, all of the variables can be considered in the research framework.

Unsurprisingly, *Store Location* is the most important variable and the conclusion is consistent with most previous literatures (Hasty and Reardon, 1997; McGlodrick, 2002; Merrilees and Miller, 1994). In addition, *Customer Visits* is also in the top five variables. The results confirm the value of physical recourse, as store location and customer visits are the performance indicators of physical recourse. The informational resource is also vital, since the variable of *Willingness to Innovate* reflects a company's ambition to know customer demands and the intention to compete in the high dynamic retail operating environment.

Customer Complaints Management demonstrates the significance of a retail company's relational resource. Finally, *Marketing Positioning* indicates the importance of a retailer's differentiation strategy. This supports findings from interviews with retail management teams. In summary, the descriptive analysis shows that all 44 selected variables can be taken into account in the research framework. The next interesting question will be: Do different retail formats, countries or departments have different viewpoints about variable importance? These issues will be discussed in the next few sections.

5.4.2 **Comparison Analysis (By Retail Format)**

The top five (most important) and bottom five (least important) variables based on the mean or median by retail formats are presented in Table 5.11. Full details are presented in Appendix G.

Table 5.11 Top Five and Bottom Five Variables (Retail Format)

Food and Beverage Stores				**Other Retail Formats**			
Top Five Variables							
Rank	Mean	Rank	Median	Rank	Mean	Rank	Median
V2	1.38	V2	1	V9	1.25	V9	1
V32	1.42	V32	1	V1	1.47	V1	1
V19	1.43	V19	1	V32	1.52	V32	1
V1	1.47	V1	1	V33	1.52	V12	1
V33	1.50	V33	1	V2	1.56	V2	1
Bottom Five Variables							
V13	2.45	V13	2	V34	2.53	V27	2.5
V43	2.45	V43	2	V23	2.55	V38	2.5
V37	2.45	V37	2.5	V42	2.56	V4	3
V35	2.71	V35	3	V35	3.00	V22	3
V22	2.85	V22	3	V22	3.16	V35	3

Most mean and median values in Table 5.11 are not greater than three. The only exception is the mean value of the *Part-Time Staff Ratio (V22)*. This exception can be explained as follows: for *non-Food and Beverage Store Formats,* changes in part-time staff ratio do not have a great impact on company performance.

Regardless of retail formats, the most important measures are: *Store Location (V2), Willingness to Innovate (V32)* and *Number of Customer Visits (V1)*, as these three variables appear in both top five variable groups. On the other hand, the least important measures are: *Joint Venture Opportunity with Competitors (V35)* and *Percentage of Part-Time Staff (V22).* These results are the same as the overall evaluation as mentioned in the previous section.

Interestingly, some measures only appear in a single retail format. For example, in *Food and Beverage Store Format*, *Market Positioning (V19)* is a significant variable, whilst it does not show identical importance in the *Other Retail Formats* group. Does this mean the viewpoints regarding the importance of variable are different between *Food and Beverage Store Format* and *Other Retail Formats*? This question needs further exploration. In this section, two nonparametric hypothesis testing techniques are employed for comparison purposes: the *Mann-Whitney U test* and the *Kolmogorov-Smirnov Test.*

The Mann-Whitney U test is the most popular nonparametric method for testing whether two independent samples are equivalent in location. Nonparametric method is used, since date type is ordinal. The hypotheses of the Mann-Whitney U test are:

H_0: The two samples are equivalent in location
H_1: The two samples are not equivalent in location

With the purpose of conducting hypothesis testing, the cases from both samples are first combined and ranked. The sum of all the ranks of each sample group is then calculated. The U statistics of each sample group can be calculated by the follow equations (Yen, 1993):

$$u_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - W_1 \tag{5.2}$$

$$u_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - W_2 \tag{5.3}$$

where u_1 and u_2 are the U statistics of the two sample groups, n_1 and n_2 are the sample size of the two sample groups, W_1 and W_2 are the sum of ranks of the two sample groups.

The final U statistics is determined as the smaller value of the u_1 and u_2 which occur. Using a confidence level of 0.05, if p-value is greater than 0.05, then the two samples are equivalent in location. In other words, if the result of the hypothesis testing is not significant, then the viewpoints from two different samples are identical. In this research, the Mann-Whitney U test is employed to compare two independent samples based on the mean value. Nevertheless, the Mann-Whitney U test is only appropriate when the number of ties between two samples is low. Given that the numbers of ties appear to be high in the samples with median, the Kolmogorov-Smirnov Test is also employed.

Kolmogorov-Smirnov (K-S) Test is used to investigate the significance of difference between two independent sample distributions. The hypothesis of the K-S test can be expressed as follow:

H_0: The two samples are from the same distribution function

H_1: The two samples are not from the same distribution function

Kanji (1999) pointed out that K-S test first determines the cumulative distribution functions of both samples. It then calculates the maximum absolute difference between two cumulative distribution functions. The basic rule is that if the maximum absolute difference is significantly large, the two distributions are considered different. Similar to the Mann-Whitney U test, a p-value is greater than 0.05 means that the two samples are from the same distribution function.

Drawing on above, the Mann-Whitney U test and K-S test are employed for evaluating whether there is existence of different viewpoints between two independent samples based on the mean value and median value respectively. The results are presented in Table 5.12.

Table 5.12 Mann-Whitney U Test and Kolmogorov-Smirnov Test (Retail Format)

Hypothesis Testing	Z Statistics	p-value
Mann-Whitney U test	-1.536	0.125
Kolmogorov-Smirnov Test	0.213	1

As the p-values of two tests are greater than 0.05, the null hypothesis of two samples coming from the same distribution function cannot be rejected. Thus, it can be concluded that there is no significant different viewpoints between the *Food and Beverage Stores Format* and *Other Retail Formats*.

5.4.3 **Comparison Analysis (By Country)**

Three countries: US, UK and Taiwan, were selected for conducting a geographical comparison analysis. In the Table 5.13, the top five and bottom five variables based on the mean or median in terms of different countries are presented. The full detailed data are also illustrated in Appendix G.

Table 5.13 Top Five and Bottom Five Variables (Country)

US				**UK**				**Taiwan**			
Top Five Variables											
Rank	Mean	Rank	Median	Rank	Mean	Rank	Median	Rank	Mean	Rank	Median
V7	1.25	V7	1	V9	1.00	V2	1	V2	1.41	V1	1
V1	1.40	V1	1	V17	1.00	V8	1	V32	1.42	V2	1
V6	1.48	V6	1	V36	1.00	V9	1	V19	1.46	V19	1
V32	1.48	V32	1	V2	1.30	V17	1	V1	1.48	V32	1
V12	1.52	V12	1	V1	1.50	V36	1	V33	1.50	V33	1
Bottom Five Variables											
Rank	Mean	Rank	Median	Rank	Mean	Rank	Median	Rank	Mean	Rank	Median
V39	2.52	V27	2.5	V7	3.00	V7	3	V43	2.43	V43	2
V27	2.60	V42	2.5	V25	3.00	V25	3	V5	2.46	V5	2
V42	2.70	V22	3	V37	3.00	V37	3	V13	2.47	V4	2.5
V22	2.85	V23	3	V38	3.00	V38	3	V35	2.73	V22	3
V23	2.95	V39	3	V22	3.10	V22	3	V22	2.91	V35	3

From Table 5.13, it is obvious that only UK views the *Part-Time Staff Ratio (V22)* as an unimportant performance measure in the retail industry, since the mean value is greater than three. In addition, regardless of the graphical area, *Store Location (V1)* appears as the most important value and the *Percentage of Part-Time Staff (V22)* shows the least important value among all variables. This conclusion is still consistent with the results in the previous sections. The next important task is to perform statistical hypothesis testing to examine whether these three countries have different viewpoints regarding the importance among variables.

Unlike the retail format comparison analysis, the number of samples is greater than two. As the Mann-Whitney U test and the Kolmogorov-Smirnov test only focus on two independent samples comparison, they are not appropriate techniques for comparisons of more than two groups. A technique to deal with this problem is the Kruskal Wallis H Test. The null hypothesis and the alternative hypothesis of the Kruskal Wallis H Test can be expressed as follows:

H_0: The K sample probability distributions are identical

H_1: At least two of the K sample probability distributions are different

The Kruskal Wallis H Test shares the same theoretical background with the Mann-Whitney U test. It assumes there are K sample groups. The Kruskal Wallis H Test first combines and ranks all cases among K samples and then computes the sum of the ranks in each sample group. Finally, the H statistics is calculated through the following formula (Kanji, 1999: 89):

$$H = \frac{12}{N(N+1)} \sum_{j=1}^{K} \frac{R_j^2}{n_j} - 3(N+1) \tag{5.4}$$

where R_j is the sum of the ranks of the j^{th} sample group, n_j is the sample size of the j^{th} sample group and N is the total sample size

The null hypothesis of equal probability distributions is rejected, if the p-value of the testing is below 0.05. The results of the Kruskal Wallis H Test based on both mean and median are expressed in Table 5.14.

Table 5.14 Kruskal Wallis H Test (Country)

Hypothesis Testing	Chi-Square Statistics	*p*-value
Kruskal Wallis H Test (by Mean)	3.565	0.17
Kruskal Wallis H Test (by Median)	2.621	0.27

The results indicate that for both mean and median values, the *p*-value of the two tests are above 0.05 and hence, the viewpoints from US, UK and Taiwan are not statistically different.

5.4.4 Comparison Analysis (By Department)

The top five and bottom five variables based on the mean or median among different departments are presented in Table 5.15.

Table 5.15 Top Five and Bottom Five Variables[1] (Department)

D1				D2				D3				D4			
Top Five Variables															
R	**Me**	**R**	**Md**	**R**	**Me**	**R**	**Md**	**R**	**Me**	**R**	**Md**	**R**	**Me**	**R**	**Md**
V19	1.40	V19	1	V32	1.22	V32	1	V1	1.26	V1	1	V1	1.27	V1	1
V32	1.40	V32	1	V33	1.22	V33	1	V2	1.26	V2	1	V2	1.45	V2	1
V33	1.50	V33	1	V19	1.33	V19	1	V32	1.36	V32	1	V9	1.60	V19	1
V1	1.60	V2	1.5	V9	1.39	V9	1	V19	1.41	V19	1	V18	1.73	V32	1
V2	1.60	V1	2	V18	1.42	V18	1	V33	1.46	V33	1	V19	1.73	V9	1.5
Bottom Five Variables															
R	**Me**	**R**	**Md**	**R**	**Me**	**R**	**Md**	**R**	**Me**	**R**	**Md**	**R**	**Me**	**R**	**Md**
V37	2.57	V4	3	V26	2.48	V26	2	V3	2.38	V3	2	V43	2.82	V43	3
V35	2.62	V17	3	V42	2.48	V42	2	V27	2.44	V27	2	V5	3.00	V5	3
V43	2.65	V22	3	V23	2.52	V23	2	V13	2.45	V13	2	V20	3.00	V20	3
V4	2.73	V35	3	V35	2.81	V35	3	V35	2.83	V35	3	V22	3.10	V22	3
V22	2.85	V43	3	V22	2.81	V22	3	V22	3.00	V22	3	V4	3.20	V4	3

D5				D6				D7			
Top Five Variables											
R	**Me**	**R**	**Md**	**R**	**Me**	**R**	**Md**	**R**	**Me**	**R**	**Md**
V7	1.00	V1	1	V2	1.32	V2	1	V30	1.33	V32	1
V1	1.52	V2	1	V32	1.53	V32	1	V18	1.44	V18	1
V2	1.52	V7	1	V1	1.63	V1	2	V19	1.44	V19	1
V32	1.58	V8	1	V19	1.63	V19	2	V32	1.44	V24	1
V33	1.58	V10	1	V44	1.63	V44	2	V1	1.56	V30	1
Bottom Five Variables											
R	**Me**	**R**	**Md**	**R**	**Me**	**R**	**Md**	**R**	**Me**	**R**	**Md**
V27	2.72	V27	3	V38	2.64	V3	3	V25	2.56	V25	2
V29	2.76	V29	3	V5	2.68	V5	3	V27	2.56	V13	3
V23	2.77	V23	3	V35	2.75	V22	3	V13	2.67	V22	3
V42	2.84	V42	3	V22	2.78	V35	3	V22	2.67	V27	3
V22	3.08	V22	3	V3	2.79	V38	3	V37	2.75	V37	3

1 Note of Table 5.15: (1) R: Rank; Me: Mean; Md: Median, (2) D1: Accounting, Finance and Auditing; D2: Operations and Store Development; D3: Investor Relations; D4: Marketing; D5: R&D and Information System; D6: Human Resources; D7: Other Departments

In addition, the frequencies of appearance of the top five and the bottom five variables among different departments are also expressed in Table 5.16. Details can be found in Appendix G.

Table 5.16 The Frequency of Variable Appearance (Department)

Top 5 by mean		**Top 5 by Median**		**Bottom 5 by Mean**		**Bottom 5 by Median**	
Variable	Frequency	Variable	Frequency	Variable	Frequency	Variable	Frequency
V1	6	V19	6	V22	7	V22	7
V19	6	V32	6	V35	4	V35	4
V32	6	V1	5	V27	3	V27	3
V2	5	V2	5	V3	2	V3	2
V33	4	V33	3	V4	2	V4	2
V18	3	V9	2	V5	2	V5	2
V9	2	V18	2	V13	2	V13	2
V7	1	V7	1	V23	2	V23	2
V30	1	V8	1	V37	2	V42	2
V44	1	V10	1	V42	2	V43	2
		V24	1	V43	2	V17	1
		V30	1	V20	1	V20	1
		V44	1	V25	1	V25	1
				V26	1	V26	1
				V29	1	V29	1
				V38	1	V37	1
						V38	1

From Tables 5.15 and 5.16, *Part-Time Staff Ratio (V22)* remains the least important performance measure. Regarding the most significant performance measures, the original top five variables: *Store Location (V2)*, *Willingness to Innovate (V32)*, *Number of Customer Visits (V1)*, *Customer Complaints Management (V33)* and *Market Positioning (V19)* are still showing their importance in different departments. Kruskal Wallis H Test is employed to explore the difference in responses from various departments (see Table 5.17).

Table 5.17 Kruskal Wallis H Test (Department)

Hypothesis Testing	Chi-Square Statistics	p-value
Kruskal Wallis H Test (by Mean)	29.975	0
Kruskal Wallis H Test (by Median)	28.647	0

As the p-values of the Kruskal Wallis H Test are below the 0.05 for both values of mean and median, the viewpoints regarding the importance of the overall 44 variables are statistically different among departments. What variables explain the existence of different points of view among departments?

The Kruskal Wallis H Test is applied again to deal with this problem. The difference between the current analysis and previous analysis is that current analysis focuses on the detection of the difference among departments in terms of each specific variable, while the previous analysis concentrates on the detection of the difference among departments in terms of the mean or median of the overall 44 variables. Therefore, the original Likert scale coding data will replace the mean or median data to be the main data source for the current analysis. The results of the Kruskal Wallis H Test are presented in Table 5.18:

Table 5.18 Kruskal Wallis H Test on Each Variable (Department)

Code	**Performance Measure**	**Department (p-value)**
V1	Number of customer visits	0.11
V2	Store location	0.29
V3	Sales of the private brand products	0.28
V4	Social responsibility	0.00
V5	Employee turnover rate	0.19
V6	Staff training	0.08
V7	Franchise system	0.13
V8	Store opening strategy	0.71
V9	Sales per store	0.36
V10	Spending-per-visit rate	0.02
V11	Internal procedures	0.00
V12	Achievement of year-end goals	0.02
V13	Investments in technology development	0.07
V14	Quality of data collection and process system	0.08
V15	Empowerment of staff	0.14
V16	Response to staff issues	0.19
V17	Inventory loss control	0.06
V18	Inventory service level	0.10
V19	Market positioning	0.08
V20	Store renovation/redecoration	0.02
V21	Expense control ability	0.15

Table 5.18 Kruskal Wallis H Test on Each Variable (Department) (Continued.)

Code	Performance Measure	Department (p-value)
V22	Percentage of part-time staff	0.62
V23	Shelf-life of new products	0.00
V24	Speed of new products development	0.03
V25	Past credit history	0.35
V26	Financial support from stockholders	1.00
V27	Internet channel development	0.56
V28	Maintaining target customer group in market diversification	0.00
V29	Following fashion trends	0.00
V30	Facing seasonal demands	0.39
V31	Openness to criticism	0.04
V32	Willingness to innovate	0.23
V33	Customer complaints management	0.07
V34	Cost sharing with suppliers for promotions	0.44
V35	Joint venture opportunity with competitors	0.87
V36	Changes in customer's preferences	0.95
V37	Changes in supplier's contract content	0.67
V38	The innovation and imitation from competitors	0.14
V39	Change of government laws	0.01
V40	Innovation of new technology equipment	0.00
V41	Change of population structure	0.30
V42	Stability of government	0.03
V43	New management system software development	0.42
V44	Change of lifestyle	0.10

From Table 5.18, it is obvious that different departments have different viewpoints in terms of 13 variables: *Social Responsibility (V4), Spending-per-visit Rate (V10), Internal Procedures (V11), Achievement of Year-end Goals (V12), Investments in Technology Development (V13), Store Renovation/Redecoration (V20), Shelf-life of New Products (V23), Speed of New Products Development (V24), Maintaining Target Customer Group in Market Diversification (V28), Following Fashion Trends (V29), Openness to Criticism (V31), Change of Government Laws (V39), Innovation of New Technology Equipment (V40)* and *Stability of Government (V42).*

The results provide two key insights. First, the original top five and bottom five variables do not belong to these 13 variables. This implies that despite the existence of dissimilar viewpoints among departments, different departments have consistent point of views in terms of both most and least important variables. The results also ensure the importance of the top five and bottom five variables.

Second, since specific departments have specific features, each is likely to differ in terms of the importance of variables. For example, the marketing department is usually more creative (or active) than other departments. Therefore, some variables may be neither relevant nor important to their routine jobs. In order to explore this issue, this research also carried out a pairwise comparison analysis by departments in terms of the 13 variables above. Since the Likert scale data is likely to create ties between the two samples, the Kolmogorov-Smirnov test is employed to perform the pairwise comparison analysis. The full detailed results are illustrated in Appendix H. In this section, an example is provided based on the pairwise comparison by *Marketing Department* and *Accounting, Finance and Auditing Department*. The results are presented in Table 5.19:

Table 5.19 Example of Pairwise Comparison Analysis

Marketing vs. Accounting, Finance and Auditing												
V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.26	0.32	0.05	0.62	0.99	1.00	0.94	0.45	1.00	1.00	0.34	0.59	1.00

Results show that *Marketing* and *Accounting, Finance and Auditing* have different point of views based on the *Internal Procedures (V11)*. It is possible to think that the jobs in the *Accounting, Finance and Auditing* are usually rule based, since listed companies need to obey the regulations in the financial market. In contrast, *Marketing* needs to have the ability to adapt to changes in market demands. Therefore, they need a more flexible working environment and heavy internal regulations will increase the operating hurdles to their jobs. This argument can be also confirmed by the importance ranking of the internal regulations for both departments—for *Accounting, Finance and Auditing Department,* the rank is 7; for *Marketing Department,* the rank is 34.

Thus far, this research has provided a descriptive analysis on variables' importance and has identified the top five and the bottom five important performance measures among 44 variables. In addition, a series of comparative analyses based on different retail formats, countries and departments were also conducted. The next section will focus on exploring the gap between the 'expected value' and the 'actual performance' among variables through a case study.

5.5 Case Study

The previous sections focused on the first part of the questionnaire in order to investigate the importance among different retail formats, countries and departments. This section will concentrate on exploring of the relationship between the first part and second part of the questionnaire. The first part of the questionnaire tries to measure the *'Expectations'* of respondents regarding the importance of each performance measure. The second part of the questionnaire requires respondents to evaluate their companies' *'Actual'* performance in terms of variables mentioned in the first part. The primary objective of this section is to explore whether there is an existence of difference between a retailer's 'expected value' and the 'actual performance' among different performance measures. Moreover, this section will also identify the case company's market position based on the R-A theory through the *'Competitive Position Matrix'*.

5.5.1 Descriptive Analysis

The case company is a Taiwanese leading chain store retailer from whom 84 responses were received. The distribution of the survey in terms of different departments is presented in Table 5.20:

Table 5.20 Frequency by Department (Case Study)

Code	Department	Frequency	Percent
D1	Accounting, Finance and Auditing	24	28.57%
D2	Operations and Store Development	17	20.24%
D3	Marketing	15	17.86%
D4	R&D and Information systems	14	16.67%
D5	Other Departments (Human Resources, Purchasing and Law)	14	16.67%
Total		84	100%

As in the Section 5.41, the department with the largest frequency is *Accounting, Finance and Auditing*. However, unlike previous sections, the *Human Resource* is combined with *Purchasing and Law* in order to conduct further statistical analysis. The overall ascending rank of mean and median values of all 44 variables in terms of both expected value and actual performance are presented in Table 5.21:

Table 5.21 Rank of Mean and Median Values (Case Study)

Expected Value (Part One of survey)				Actual Performance (Part Two of survey)			
Rank	Mean	Rank	Median	Rank	Mean	Rank	Median
V19	1.48	V1	1.00	V25	1.67	V2	2.00
V2	1.49	V2	1.00	V4	1.93	V4	2.00
V1	1.52	V19	1.00	V32	2.02	V11	2.00
V32	1.55	V32	1.00	V36	2.02	V12	2.00
V33	1.62	V33	1.50	V26	2.03	V17	2.00
V9	1.69	V6	2.00	V44	2.15	V18	2.00
V7	1.71	V7	2.00	V30	2.24	V19	2.00
V8	1.73	V8	2.00	V12	2.27	V20	2.00
V18	1.73	V9	2.00	V2	2.30	V21	2.00
V30	1.76	V10	2.00	V29	2.32	V24	2.00
V44	1.82	V11	2.00	V19	2.33	V25	2.00
V29	1.86	V12	2.00	V18	2.35	V26	2.00
V36	1.86	V14	2.00	V24	2.37	V29	2.00
V6	1.87	V15	2.00	V33	2.37	V30	2.00
V10	1.88	V16	2.00	V11	2.39	V32	2.00
V14	1.98	V17	2.00	V17	2.40	V33	2.00
V40	1.99	V18	2.00	V21	2.41	V36	2.00
V24	2.00	V20	2.00	V1	2.44	V41	2.00
V28	2.00	V21	2.00	V9	2.46	V44	2.00
V21	2.10	V23	2.00	V20	2.46	V1	2.50
V41	2.11	V24	2.00	V28	2.48	V7	2.50
V23	2.11	V26	2.00	V41	2.49	V28	2.50
V39	2.12	V28	2.00	V8	2.49	V3	3.00
V12	2.14	V29	2.00	V7	2.50	V5	3.00
V11	2.18	V30	2.00	V31	2.52	V6	3.00
V17	2.19	V31	2.00	V40	2.53	V8	3.00
V31	2.20	V34	2.00	V23	2.55	V9	3.00
V38	2.20	V36	2.00	V10	2.58	V10	3.00
V16	2.22	V37	2.00	V38	2.60	V13	3.00
V15	2.23	V38	2.00	V34	2.61	V14	3.00
V25	2.35	V39	2.00	V13	2.63	V15	3.00
V34	2.36	V40	2.00	V39	2.65	V16	3.00
V4	2.37	V41	2.00	V14	2.65	V22	3.00
V26	2.39	V42	2.00	V42	2.74	V23	3.00
V42	2.39	V44	2.00	V22	2.75	V27	3.00
V37	2.42	V4	2.50	V15	2.79	V31	3.00
V13	2.52	V25	2.50	V27	2.80	V34	3.00
V43	2.54	V3	3.00	V37	2.87	V35	3.00

Table 5.21 Rank of Mean and Median Values (Case Study) (Continued.)

Expected Value (Part One of survey)				Actual Performance (Part Two of survey)			
V20	2.55	V5	3.00	V5	2.89	V37	3.00
V27	2.56	V13	3.00	V3	2.89	V38	3.00
V3	2.60	V22	3.00	V43	2.94	V39	3.00
V5	2.62	V27	3.00	V6	2.95	V40	3.00
V35	2.73	V35	3.00	V16	3.05	V42	3.00
V22	2.96	V43	3.00	V35	3.15	V43	3.00

Regarding the expected value among variables, the top five important variables are still the same as previous sections. However, the bottom five variables are slightly different. Particularly, the variable of *Changes in Supplier's Contract Content (V37)* does not appear in the bottom five variables in this case company. This measure is not in the bottom five variables mainly because of this company has strong and long term relationship with its suppliers. Any change of the contract content will have great impacts on this company's performance, since this company is difficult to find substitute suppliers in the short term.

With regards to actual performance among variables, two variables: *Joint Venture Opportunity with Competitors (V35)* and *Response to Staff Issues (V26)* were weak, since their average values are greater than three. Moreover, there is only one variable from the actual performance—*Willingness to Innovate (V32)*—in the group of top five important variables. Thus, a difference between 'expected value' and the 'actual performance' was suspected. In fact, for most variables, the actual performance was not in line with the expected importance. The case company's mean and median values based on the actual performance are usually higher than the mean and median values based on the expected importance. In other words, the gap between the 'expected value' and the 'actual performance' among variable is noticeable. Is this gap statistically significant? This issue will be discussed in the next section.

5.5.2 Comparative Analysis (By Department)

Unlike previous statistical analysis, this section employed the *Wilcoxon Signed Ranks Test* be the main hypothesis testing technique. The main consideration is that the two samples in this section are related, that is, from the same respondent. This is in contrast to statistical analyses carried out the previous sections, where the samples are independent, that is, from different respondents. Wilcoxon Signed Ranks test compares the distributions of two related variables. The null hypothesis and the alternative hypothesis of Wilcoxon Signed Ranks test can be presented as follows:

H_0: The population median of the paired differences of the two samples is 0

H_1: The population median of the paired differences of the two samples is not 0

Wilcoxon Signed Ranks test first calculates the absolute difference $D = |x_a - x_b|$ for each pair sample. If $D = 0$, then the sample should be omitted. It then ranks the pair sample based on the D values and assigns a sign '+', if $x_a - x_b > 0$ or a sign '–', *if* $x_a - x_b < 0$. The next step is to calculate the sum of ranks based on the '+' sign samples (S^+) and '–' sign samples (S^-). The final choice of the sum of ranks depends on the lower value between the S^+ and S^-. If the sample size is large enough, the Wilcoxon Signed Ranks test is approximately a normal distribution. Hence, the hypothesis testing can be performed by calculating the Z statistics:

$$Z = \frac{S - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \tag{5.5}$$

where $S = \min\{S^+, S^-\}$ and n is the number of total samples

If the p-value is below 0.05, the null hypothesis will be rejected and it means that the gap between the 'expected value' and the 'actual performance' is statistically significant. The results of the Wilcoxon Signed Ranks test in terms of both mean and median values are presented in Table 5.22:

Table 5.22 Wilcoxon Signed Ranks Test Results (Mean and Median Data)

Hypothesis Testing	Z Statistics	p-value
Wilcoxon Signed Ranks Test (by Mean)	-4.849	0
Wilcoxon Signed Ranks Test (by Median)	-4.339	0

From Table 5.22, it is clear that despite the mean or the median, the *p*-value of the Wilcoxon Signed Ranks test is below 0.05. It can be concluded that the gap between the 'expected value' and the 'actual performance' is statistically significant among variables. Which variables explain the existence of the gap between the 'expected value' and the 'actual performance'? Again, the Wilcoxon Signed Ranks test is applied to the original Likert scale data in order to investigate this question. The results are illustrated in Table 5.23:

Table 5.23 Wilcoxon Signed Ranks Test Results (Likert Scale Data)

Code	Performance Measure	*p*-value
V1	Number of customer visits	0.00
V2	Store location	0.00
V3	Sales of the private brand products	0.01
V4	Social responsibility	0.00
V5	Employee turnover rate	0.04
V6	Staff training	0.00
V7	Franchise system	0.00
V8	Store opening strategy	0.00
V9	Sales per store	0.00
V10	Spending-per-visit rate	0.00
V11	Internal procedures	0.06
V12	Achievement of year-end goals	0.10
V13	Investments in technology development	0.43
V14	Quality of data collection and process system	0.00
V15	Empowerment of staff	0.00
V16	Response to staff issues	0.00
V17	Inventory loss control	0.04
V18	Inventory service level	0.00
V19	Market positioning	0.00
V20	Store renovation/redecoration	0.38
V21	Expense control ability	0.00
V22	Percentage of part-time staff	0.08
V23	Shelf-life of new products	0.00
V24	Speed of new products development	0.00
V25	Past credit history	0.00
V26	Financial support from stockholders	0.00
V27	Internet channel development	0.03
V28	Maintaining target customer group in market diversification	0.00
V29	Following fashion trends	0.00
V30	Facing seasonal demands	0.00
V31	Openness to criticism	0.00

Table 5.23 Wilcoxon Signed Ranks Test Results (Likert Scale Data) (Continued.)

Code	Performance Measure	*p*-value
V32	Willingness to innovate	0.00
V33	Customer complaints management	0.00
V34	Cost sharing with suppliers for promotions	0.00
V35	Joint venture opportunity with competitors	0.00
V36	Changes in customer's preferences	0.04
V37	Changes in supplier's contract content	0.00
V38	The innovation and imitation from competitors	0.00
V39	Change of government laws	0.00
V40	Innovation of new technology equipment	0.00
V41	Change of population structure	0.00
V42	Stability of government	0.00
V43	New management system software development	0.00
V44	Change of lifestyle	0.00

The results indicate that the gaps between the 'expected value' and the 'actual performance' are statistically significant among almost all the variables. The only exceptions are: *Internal Procedures (V11)*, *Achievement of Year-end Goals (V12)*, *Investments in Technology Development (V13)*, *Store Renovation/Redecoration (V20)* and *Percentage of Part-time Staff (V22)*. Another interesting question will be: Is this situation the same among different departments?

In order to explore this issue, this research first calculates the mean and median values in terms of both parts of the questionnaire among different departments. It then carried out a comparison analysis by using the Wilcoxon Signed Ranks test based on both mean and median data. The results are shown in Table 5.24 and 5.25 respectively. The original mean and median data among departments are presented in Appendix I.

Table 5.24 Wilcoxon Signed Ranks Test Results by Department (Mean)

Department	**Z-statistics**	***p-value***
MARKETING	-5.7889	0.0000
OPERATIONS and STORE DEVELOPMENT	-5.7252	0.0000
ACCOUNTING, FINANCE and AUDITING	-5.7791	0.0000
R&D and INFORMATION SYSTEM	-5.7007	0.0000
OTHER DEPARTMENTS	-5.7880	0.0000

Table 5.25 Wilcoxon Signed Ranks Test Results by Department (Median)

Department	**Z-statistics**	***p-value***
MARKETING	-4.6252	0.0000
OPERATIONS and STORE DEVELOPMENT	-4.6127	0.0000
ACCOUNTING, FINANCE and AUDITING	-4.9630	0.0000
R&D and INFORMATION SYSTEM	-3.4157	0.0006
OTHER DEPARTMENTS	-3.8517	0.0001

Clearly, the results show that for both mean and median, the viewpoints between the 'expected value' and the 'actual performance' are statistically different among various departments.

Drawing on the results in this section, almost all the variables that were considered important performance measures were the ones where the respondents believed the company did not perform well on. Does it mean this case company simply does not perform well? In fact, this sample company's policy has to set a higher standard for their annual objectives in order to encourage employees to face challenges. Even if they have good performance in terms of some specific issues, they will still think they can do better. Therefore, the results tend to be conservative based on the actual performance in this case company are as a matter of course.

The survey data in this case study can also be employed to determine the market position of the case company based on the R-A theory. Furthermore, the core resources of the sample company can also be defined. The next section will discuss the relationship between the survey results and the application of R-A theory.

5.5.3 Market Position Determination of the Case Company

Chapter Three introduced the *Competitive Position Matrix*, which is used to define a company's market position based on the R-A theory (see Table 3.3). The criteria to identify a company's competitive position are: *Relative Resource Costs* and *Relative Resource-Produced Value*. It has been argued that if a company can occupy the three cells located at the upper right-hand side in the competitive position matrix, it will achieve superior comparative advantage in the marketplace, and achieve the goal of superior financial performance.

There is a strong connection between the competitive position matrix and the questionnaire design in this research. On the one hand, the first part of the questionnaire is used to evaluate the importance of different resources. The importance of a specific resource can be regarded as the relative costs of the resource, since the importance of a variable reflects its expected value. On the other hand, the second part of the survey is employed to assess a company's actual performance in terms of the variables mentioned in the first part. In other words, the second part of the survey tries to obtain the insights regarding the value produced from the resources.

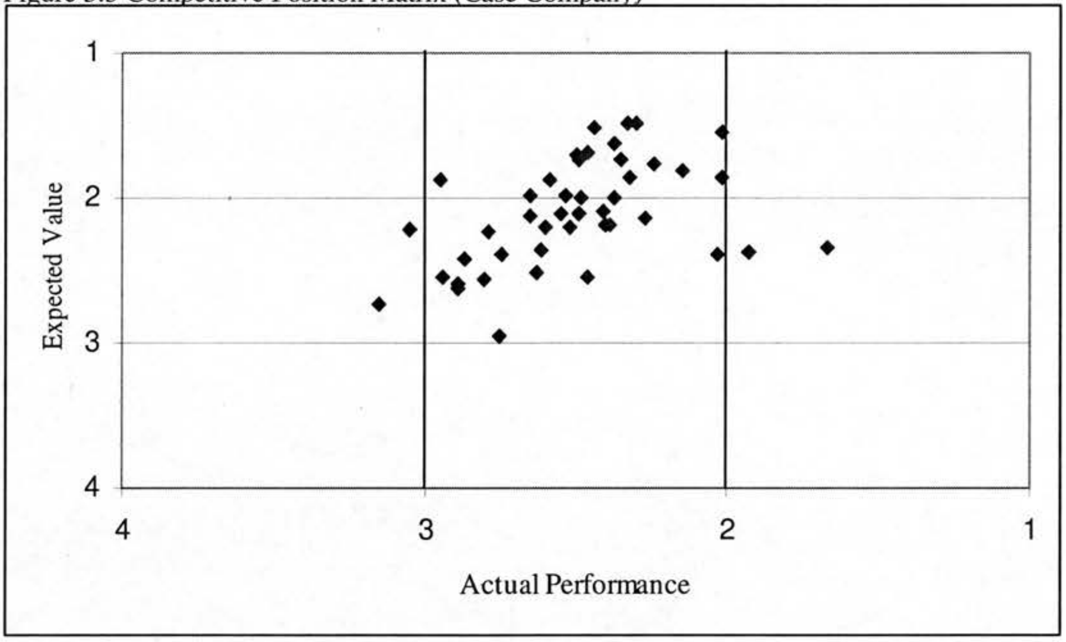
Therefore, based on the groundwork of the competitive position matrix, the survey data can be used to identify the case company's market position. Prior to explore this issue, it is useful to redefine the competitive position matrix and it is presented in Table 5.26.

Table 5.26 Competitive Position Matrix for Survey Analysis

		Actual Performance of Resource		
		Lower	Parity	Superior
Expected Importance of Resources	Higher	1 Indeterminate Position	2 Competitive Advantage	3 Competitive Advantage
	Parity	4 Competitive Disadvantage	5 Parity Position	6 Competitive Advantage
	Lower	7 Competitive Disadvantage	8 Competitive Disadvantage	9 Indeterminate Position

Similarly, if a company can occupy one of three cells (cell 3, cell 2 or cell 6) in this new competitive position matrix, then it will have comparative advantage in the marketplace. Clearly, cell 3 is the perfect situation, since it means that firms show superior performance based on the most important variables. With regards to the two other cells, although they are not as good as the cell 3, they still show their potential value to obtain the competitive advantage in the market. Therefore, if the case company's resources are in these three cells, these resources can be viewed as the strengths or the core resources of the case company. The mean data in Table 5.21 is selected to explore this issue and the results are presented in the Figure 5.3.

Figure 5.3 Competitive Position Matrix (Case Company)



From Figure 5.3, it is clear that the case company does not have any variables located in cell 3. However, it has 19 variables in cell 2 and two variables in cell 6. All the variables with higher importance, that is with the mean values between 1 and 2, show at least parity actual performance. Although this company does not have any core resource in cell 3, it is still a leading retailer in the Taiwanese market, since the variables in cell 2 and cell 6 can be viewed as the case company's competitive strengths. The core resources in cell 2 and cell 6 are listed in Table 5.27:

Table 5.27 Core Resources of Case Company

Cell 2: Higher Expected Importance and Parity Actual Performance	
Code	Core Resources
V1	Number of customer visits
V2	Store location
V6	Staff training
V7	Franchise system
V8	Store opening strategy
V9	Sales per store
V10	Spending-per-visit rate
V14	Quality of data collection and process system
V18	Inventory service level
V19	Market positioning
V24	Speed of new products development
V28	Maintaining target customer group in market diversification
V29	Following fashion trends
V30	Facing seasonal demands
V32	Willingness to innovate
V33	Customer complaints management
V36	Changes in customer's preferences
V40	Innovation of new technology equipment
V44	Change of lifestyle
Cell 6: Parity Expected Importance and Superior Actual Performance	
Code	Core Resources
V4	Social responsibility
V25	Past credit history

5.6 Concluding Remarks

The primary objective of this chapter was to introduce a research survey in order to obtain more information from context on the initial research framework. This chapter begins by introducing the key issues related to the survey design, including: sampling strategy, variable selection, questionnaire format and survey conducting approach.

Regarding the sampling strategy, this research adapted the *Two-Stage Geographical Cluster Sampling* method with the aim of conducting a comparative analysis in terms of three different geographical areas: North America, Europe and Asia. In order to select appropriate performance measures for questionnaire design, this research regrouped all variables into: *'quantifiable measure and available data'* group, *'quantifiable measure but no available data'* group, and *'difficult to quantify'* group. The variables in the *'quantifiable measure but no available data'* group, and *'difficult to quantify'* group will be the potential candidate variables in the survey, since these variables require further study to quantify them in order to achieve the objective of examining the importance among variables. Overall, 44 variables were selected based on the R-A theory for final questionnaire design.

There are two primary sections in the questionnaire. The first section focuses on measuring the *expectations* from respondents regarding the importance of each key performance measure. The second section requires respondents to evaluate their companies' *'Actual'* performance in terms of variables mentioned in the first section. All these two sections are designed by using a Likert scale in order to evaluate the different levels regarding the 'expected value' and 'actual performance' among all 44 variables.

The research used e-questionnaire as the main survey instrument, due to the assumptions of time and cost savings as well as convenience to respondents in most cases. However, the primary drawback of the e-questionnaire is low responses rate, since respondents usually delete emails with on-line survey. With the aim of overcoming this drawback, not only were the e-questionnaire sent to the sample companies, but also each sample company was contacted by phone in order to request them to complete the e-questionnaire.

Prior to performing the formal survey, a pilot survey was carried out in order to ensure the quality of the questionnaire. Some reflections were made based on the response rate, validity and reliability of the pilot questionnaire. For example, this research refined some unclear terms with the purpose of ensuring every variable could measure the concept appropriately. Cronbach's alpha was employed to assess the reliability of the pilot survey and the results indicated that the reliability of the pilot survey was acceptable.

Finally, 435 e-questionnaires were sent out to US, UK and Taiwan and 151 responses were received. The overall response rate was 34.71%. This research first conducted a descriptive analysis based on the mean and median values among all variables and the results showed that all of the 44 variables could be considered in the research framework. The primary reason is that the importance of all the variables is above the level of *'Moderately Important'*. In addition, regardless of mean and median values, the top five and bottom five important variables are identical.

With regards to the comparative analysis in terms of different retail formats, countries and departments, three nonparametric hypothesis testing techniques were employed: *Mann-Whitney U test*, *Kolmogorov-Smirnov Test* and *Kruskal Wallis H Test*. The results indicated that there is no difference regarding variable importance among different retail formats and countries. However, the viewpoints regarding the importance of the overall 44 variables are statistically different among departments. The main reason of the dissimilarity is that as different departments have different concerns, so are highly likely to have different viewpoints in terms of some specific variables.

For example, *Marketing Department* and *Accounting, Finance and Auditing Department* have different point of views based on the variable of *Internal Procedures (V11)* possibly because the jobs in the *Accounting, Finance and Auditing Department* are usually rule based but those in *Marketing* require the ability to quick face the various market demands. Therefore, they need a more flexible working environment.

Apart from the analysis of the importance among variables, this chapter also carried out a case study in order to investigate the difference between the 'expected value' and the 'actual performance' among variables. The *Wilcoxon Signed Ranks Test* was selected to be the main nonparametric hypothesis testing technique to explore this issue. The results showed that the viewpoints between the 'expected value' and the 'actual performance' are statistically different overall. The primary reason is that this case company's policy is to set a higher standard for their annual objectives in order to encourage employees to face challenges. Even if they have good performance in terms of some specific variables, they will still think they can do better. Therefore, the results tend to be conservative based on the actual performance as a matter of course.

The data in this survey is also useful for identifying the case company's market position based on the R-A theory. By using the competitive position matrix, the company's 21 core resources were identified. These core resources facilitate to explain the case company's competitive strengths in the marketplace and to determine its market position.

Thus far, it can be concluded that the results of the robust examination of the research framework were satisfactory, since all the 44 variables could be taken into account in the framework. However, due to the limitation of the length of survey, this research only considered 44 variables to be the main performance measures in the questionnaire. The survey could have been extended by considering more variables.

The contents of Chapters Four and Five are related to framework development. Chapters Six to Seven, the focus will place on the research model construction. Key issues regarding data collection and variable selection will be introduced. In addition, an illustration of the key variables selection procedures, such as principal component analysis and stepwise regression approach will also be presented.

Part Four

Default Prediction Model Construction

- Chapter Six: ***Model Construction: Data Collection and Key Variables Determination***
- Chapter Seven: ***Model Construction: Modelling Techniques and Cross-Validation Process***

Chapter Six reviews the sample selection criteria, the data arrangement as well as the selection procedures for key variables, such as principal component analysis and stepwise regression approach. The selected key variables and principal components will also be described. *Chapter Seven* then concentrates on the modelling procedure, including an introduction of the selected credit scoring techniques and cross-validation approaches.

Chapter SIX

Model Construction: Data Collection and Key Variables Determination

6.1 Introduction

In Chapters Four and Five, the focus was on the development of a research framework based on the R-A theory of competition. This chapter and the next will concentrate on the key points relating to default prediction model construction.

First, quantitative data collection issues concerning variable selection, variable regrouping, sample collection criteria and sources of data collection are discussed. Second, the process of determining key variables is explained. In addition, the techniques of key variables determination will also be introduced. Finally, findings from the previous sections are summarized.

6.2 Data Collection

6.2.1 Variable Selection

As mentioned in Chapter Four, 170 variables were found based on the framework of R-A theory. All 170 variables in the research framework can be grouped under: *'quantifiable measure and available data'*, *'quantifiable measure but no available data'*, and *'difficult to quantify data'*. Out of the 170 variables, only the 67 variables in the *'quantifiable measure and available data'* group were selected for final model development, since these variables could be easily calculated and collected from financial databases or annual reports.

The researcher also explored the use of qualitative variables (103 variables) in the assessment of business failure. The development of a survey instrument served to collect appropriate information on variables in the *'quantifiable measure but no available data'* and *'difficult to quantify'* groups from retail management in the USA. The incorporation of such information into the default prediction model could be contemplated using *Bayesian* techniques. However, it was extremely difficult to obtain qualitative information from already distressed companies, since most distressed companies have bankrupted already and disappeared from the market. Although some distressed firms still existed in the market, they do not reveal any information to outside users with the purpose of avoiding hurdles for financial reconstruction. An option would be to explore companies that were deemed *'close'* to distress. Nevertheless, information may also be hard to get hold of. Due to such difficulties, work on qualitative variables as a means of predicting financial distress is still at a preliminary stage. The thesis will only concentrate on the variables in the *'quantifiable measure and available data'* group for retail default prediction model construction. The list of the 67 variables is presented in Table 6.1:

Table 6.1 Variable List

Code	Variable
V1	EBIT margin (%)
V2	EBITDA margin (%)
V3	EBITDAR margin (%)
V4	Pre-tax profit margin (%)
V5	Pre-tax profit / capital employed
V6	Net profit margin (%)
V7	Gross margin (%)
V8	SG&A as % of net sales (%)
V9	EBIT on capital
V10	Return of total assets (%)
V11	Return on total equity (%)
V12	Operation margin (%)
V13	Dividend payout ratio (%)
V14	Inventory turnover
V15	Receivable turnover
V16	Payables turnover
V17	Total assets turnover
V18	Fixed assets turnover
V19	Current ratio
V20	Acid ratio
V21	Cash ratio
V22	Net operating cash flow / capital expenditures
V23	Cash dividend cover
V24	Fixed charge cover
V25	Interest cover
V26	Net operating cash flow / total debt

Table 6.1 Variable List (Continued.)

Code	Variable
V27	EBITDA / interest
V28	Total debt / discretionary cash flow
V29	Debt ratio
V30	Debt / EBITDA
V31	Leased-adjusted net debt / EBITDAR
V32	Net debt as % of market capitalisation
V33	Total debt / (total debt + market value of equity)
V34	Debt to equity ratio
V35	P/E ratio
V36	Net sales (log)
V37	Total assets (log)
V38	Market share by retail sector (based on sales) (%)
V39	Market share by retail sector (based on gross margin) (%)
V40	Total capital (log)
V41	Number of payrolls (log)
V42	Operation cash flow (log)
V43	Store numbers (log)
V44	Like-for-like sales growth (%)
V45	Market value growth (%)
V46	Total capital growth (%)
V47	Number of payrolls growth (%)
V48	EBIT growth (%)
V49	Number of stores growth (%)
V50	The operating income growth (%)
V51	Market capitalisation / net assets
V52	Sales per employee
V53	EBIT per employee
V54	Net cash cycle
V55	Main market sales as percentage of total sales (%)
V56	The five-year correlation coefficient between GDP and total sales
V57	The five-year correlation coefficient between average interest rate and total sales
V58	The five-year correlation coefficient between unemployment rate and total sales
V59	The five-year correlation coefficient between disposable income and total sales
V60	The five-year correlation coefficient between birth rate and total sales
V61	The five-year correlation coefficient between death rate and total sales
V62	The five-year correlation coefficient between age structure ratio (0-14 years old) and total sales
V63	The five-year correlation coefficient between age structure ratio (65 years and over) and total sales
V64	The five-year correlation coefficient between total government spending for R&D and total sales
V65	The five-year correlation coefficient between government debt / GDP and total sales
V66	The five-year correlation coefficient between government avenue / GDP and total sales
V67	The five-year correlation coefficient between government expense / GDP and total sales

6.2.2 Variable Regrouping

As mentioned in Chapter Two, a number of previous studies argued that the external environmental factors have great impacts on the accuracy of a bankruptcy prediction model. Taking this into consideration the 67 variables are regrouped into two categories: *Internal Resources Group (G1)* and *External Factors Group (G2).*

Variables (from V1 to V55) in the *Internal Resources Group (G1)* are mainly related to financial accounting theory and form the basis of most previous studies on default prediction. Moreover, certain measures are specific to the operational side of the retail industry, such as the reach ability and the brand strength. *External Factors Group (G2)* consists of 12 variables (from V56 to V67) calculated as the five-year correlation coefficients between external environmental factors (such as GDP, interest rate or unemployment rate) and the total sales in each sample company. In order to detect the influences from the external environment, all the performance measures in G1 and G2 can be further regrouped into G1 and G12 (where G12 is G1 plus G2). It can be argued that if G12 has better performance than G1, then influences from the external environment are significantly affect the performance of a default prediction model.

The analysis did not use the variables over time or coarse classification, but rather, and exploratory approach was adapted using fast testing variables. Further work could use these variables in model. Traditionally, default prediction studies have used linear measures. This research could have done coarse classification to become more non-linear in form. Some of the approaches used are non-linear by nature (such as, SMO, Neural Network and Recursive Partitioning). If these approaches performed better with large positive differences, then non-linear variables would have been picked up in these approaches. However, this was not the case. Hence, it is not certain that the variables over time or coarse classification would be necessary to employ in model.

6.2.3 Sample Selection Criteria

6.2.3.1 Sample Selection Criteria for Non-defaulting Companies

In connection with the sample selection of non-defaulting companies, five criteria were considered. Only publicly listed companies were chosen. Given that listed companies had to abide by regulations in the financial market, their financial information tended to be more open and transparent than that of private companies.

In addition, small companies were included based on SBA size standards[1]. This is an enhancement from previous studies using *Moody's Industrial Manual* and the *Compustat database*. These data sources only include newsworthy companies and are likely to exclude small companies despite the fact that small companies are more likely to face financial distress.

Although Knott and Posen (2005) argued that new firms had great probability to face financial distress and should be considered in any bankruptcy prediction model, the present study only considered those public sample companies that had been listed for at least three years. One reason for omitting newly listed companies is that a number of studies show that newly listed stocks have abnormal returns after the public announcements of listing (Sanger and McConnell, 1986). In order to avoid the influences from newly listed companies, especially for those market relevant measures, no healthy company listed after December 2000 was to be included. In addition, it should be borne in mind that newly listed is not the same as new.

Furthermore, it must be noted that e-retailers are not considered because their performance measures are different from those of traditional retailers. Finally, even if a sample company satisfied the previous four criteria[2], it was excluded if its data were not complete. As a result of applying the five criteria above, 67 different retail performance measures were collected from a dataset of 195 non-defaulting US retail companies over the time period of 1998 to 2002.

6.2.3.2 Sample Selection Criteria for Defaulting Companies

What is the definition of financial distress? Altman (1983) introduced a definition of financial distress based on an accounting and finance point of view (also see Ross et al., 1999). He pointed out that the definition of financial distress had two themes: stock-based insolvency and flow based insolvency. Stock-based insolvency occurs

---

1 SBA's size standards define whether a business entity is small and, thus, eligible for government programs and preferences reserved for "small business" concerns. Size standards have been established for types of economic activity, or industry, generally under the North American Industry Classification System (NAICS). Information available at: http://ecfr.gpoaccess.gov/

2 Publicly listed, small size consideration, no newly listed companies, and no e-retailers

when a company's total liabilities are greater than its total assets, while flow-based insolvency occurs when a company's operating cash flow cannot meet its routine obligations.

In this research, the definition of financial distress was founded on a legal viewpoint. Based on the US federal bankruptcy law[3], a financially distressed company may use the bankruptcy code of *Chapter 11* to reorganize its financial structure and try to recover from distress, or that of *Chapter 7*, to go into liquidation and stop all business operations. Drawing on this insight, any company filing for the bankruptcy code of Chapter 11 or Chapter 7, were deemed to be under financial distress and selected for the research. The USA retail sector was chosen because of the clear definitions and reporting of financial distress through Chapters 7 and 11.

An important issue is the timing of distressed firms' data. As mentioned in Chapter Two, Ohlson (1980) pointed out that the most recent financial statement prior to default for a distressed firm may only be available after bankruptcy is filed—if this company files for bankruptcy between the date of financial statements releasing and the date of the financial year. Ohlson suggested that for these companies, the financial statements prior to the financial distress year should be viewed as the *last* report, since reports after financial distress would usually include the adjustments from auditors in light of the bankruptcy filing. Considering Ohlson's (1980) viewpoints, the researcher employed an even more conservative approach to deal with this problem—all defaulting companies' data prior to the financial distress year was considered as the last report. Overall, data were collected from 51 financially distressed firms and these companies were divided into five groups (from group A to group E) in terms of five different time scales (see Table 6.2).

Table 6.2 Descriptions of Time Scales of Distressed Firms' Data

Group	Number of Failed Firms	Financial Distress Year	Data Collection Time Scale
A	5	2003	From 1998 to 2002
B	13	2002	From 1997 to 2001
C	15	2001	From 1996 to 2000
D	12	2000	From 1995 to 1999
E	6	1999	From 1994 to 1998
Total	51		

3 Information is available at: http://www.uscourts.gov/bankruptcycourts.html

6.2.4 Data Sources

Due to a large range of variables in this research, it is not possible to collect data from only single data sources. In this research, data were collected from four main sources: 1) *Accounting and Finance Databases*, such as, DataStream and OSIRIS, 2) *Annual Report from Each Sample Company*, 3) *Government Publications*, such as, Budget of the United States Government and 4) *Other Sources*, such as documents from *Organisation for Economic Co-operation and Development (OECD).* For example, financial statements usually do not reveal the information about a retailer's store number. Hence, it is necessary to collect information on the store number by studying each company's website or annual report.

Drawing on the above, a dataset of 195 US healthy retailers and 51 US distressed retailers with 67 variables over five time periods: 1994-1998, 1995-1999, 1996-2000, 1997-2001 and 1998-2002 were collected. Yet obviously with such a large number of variables to choose from there is a danger of over-fitting and so there is a need to reduce the number of variables and to determine the key variables for final default prediction model construction. This issue will be discussed in the next section.

6.3 Key Variables Determination

In Chapter Two, it was mentioned that due to the lack of theoretical groundwork for variable selection, most financial distress prediction studies took into account a large number of variables in order to consider all potential useful variables. The procedure of variable reduction and key variables determination was necessary to avoid the situation whereby too many variables in a model would tend to overfit the model and hence make poor predictions.

Principal Components Analysis (PCA) and *Stepwise Regression Approach* were the most two popular methods for variable reduction and selection in previous bankruptcy prediction studies. This research will also adapt these two approaches to determine key variables, or, principal components. Prior to selecting the final variables, some key issues include: time-scale consideration, outlier elimination and univariate analysis, were carried out in order to ensure the quality of key variables or principal components. These are discussed below.

6.3.1 Time Scale Consideration

If a performance measure was *'truly'* important, then it would be significant in different time periods. This study considers five different model time scales with the purpose of determining the key variables and which are presented in Table 6.3:

Table 6.3 Model Time Scales

Code	Healthy Firms Time Scale	Distressed Firms Time Scale[4]
M1	1998~2002 five years average data	• Group A financial distress firms: 2002 data • Group B financial distress firms: 2001 data • Group C financial distress firms: 2000 data • Group D financial distress firms: 1999 data • Group E financial distress firms: 1998 data
M2	1998~2002 five years average data	• Group A financial distress firms: 2001 data • Group B financial distress firms: 2000 data • Group C financial distress firms: 1999 data • Group D financial distress firms: 1998 data • Group E financial distress firms: 1997 data
M3	1998~2002 five years average data	• Group A financial distress firms: 2000 data • Group B financial distress firms: 1999 data • Group C financial distress firms: 1998 data • Group D financial distress firms: 1997 data • Group E financial distress firms: 1996 data
M4	1998~2002 five years average data	• Group A financial distress firms: 1999 data • Group B financial distress firms: 1998 data • Group C financial distress firms: 1997 data • Group D financial distress firms: 1996 data • Group E financial distress firms: 1995 data
M5	1998~2002 five years average data	• Group A financial distress firms: 1998 data • Group B financial distress firms: 1997 data • Group C financial distress firms: 1996 data • Group D financial distress firms: 1995 data • Group E financial distress firms: 1994 data

4 see Table 6.2

In Table 6.3, model 1 (M1) evaluates the model prediction performance *one year* before financial distress. M2, M3, M4 and M5 are designed for assessing the model prediction utility two, three, four and five years before financial distress respectively. Moreover, an initial interest of this research was the timescale effect, whether one should use data just prior to the potential financial distress or some time before. Hence, a series of models were fitted to M1 to M5 to enable an evaluation of prediction performance from one to five years before financial distress. The results of the five-year time scales comparative analysis will be introduced in Chapter Eight.

6.3.2 Outlier Elimination

The K-means cluster analysis is applied to remove outlier impacts. Cluster analysis serves to group objects based on the characteristics they possess (Hair Jr. et al., 1998) that is, according to their similarity (Fielding and Gilbert, 2000). The basic rules for grouping are: minimize the variability within clusters and maximize the variability between clusters (Chou, 2002). Based on the criterion of similarity, it is expected that most objects will be grouped in one cluster and the rest of the objects will be grouped to other clusters. Objects that cannot be grouped due to lack of similarity may be viewed as outliers.

A 10-means cluster analysis was selected to remove outliers, using 5% as the outlier elimination rate. New sample composition after the 10-means cluster analysis provided an elimination rate between 4.47% and 5.69% among five different time-scale models. Table 6.4 shows the new sample composition after the process of outlier elimination.

Table 6.4 New Sample Composition After 10-means Cluster Analysis

Model by time scales	Healthy Firms	Distressed Firms	Total Firms	Elimination Rate
M1	186	46	232	5.69%
M2	186	47	233	5.28%
M3	187	48	235	4.47%
M4	185	47	232	5.69%
M5	184	48	232	5.69%

6.3.3 Logistic Univariate Analysis

According to Hosmer and Lemeshow (2000), the variable selection process should begin with a univariate analysis of each independent variable. An independent variable should be used in further multivariate analysis only if the *p*-value of an independent variable's univariate test is below 0.25. The main reason for selecting the significance level of 0.25 is that some variables may not be significant at the traditional significance level of 0.05, but when taken together with other variables may prove to be significant. In order to include all potential variables, a significance level of 0.25 is deemed to be suitable for logistic univariate analysis.

Drawing on above, potential candidate variables were selected based on a series of logistic univariate analyses with the significance level of 0.25. Those variables that were eliminated during the process are presented in Table 6.5:

Table 6.5 Eliminated Variables after Logistic Univariate Analysis

Model	Variables with *p*-value >= 0.25
M1	V14, V16, V17, V18, V28, V30, V34, V40, V46, V48, V50, V52, V54
M2	V14, V16, V17, V18, V19, V28, V30, V31, V44, V46, V47, V49, V50, V52, V54, V55
M3	V7, V9, V13, V14, V16, V17, V19, V21, V28, V30, V35, V44, V45, V46, V47, V48, V49, V50, V54, V55
M4	V7, V9, V13, V14, V16, V17, V18, V19, V20, V21, V28, V30, V31, V34, V35, V40, V44, V46, V47, V48, V49, V50, V55, V60
M5	V5, V7, V9, V11, V13, V14, V16, V17, V18, V19, V28, V30, V31, V34, V35, V40, V45, V49, V50, V54, V60, V61

6.3.4 Principal Components Analysis

Principal Components Analysis (PCA) can be used to reduce the number of variables by transforming the original variables into a new set of linear combinations—called *Principal Components*. Stevens (1992) pointed out that PCA first finds the linear combination of variables, which accounts for the maximum of variance, to be the first principal component (PC_1).

$$PC_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \dots a_{1n}x_n \tag{6.1}$$

The variance of PC_1 is the largest eigenvalue in the sample covariance matrix. The coefficients ($a_{11}, a_{12} \dots a_{1n}$) are called *eigenvectors.* The sum of the square of all eigenvectors is equal to the eigenvalue of PC_1. The second principal component (PC_2) is found by accounting for the second largest amount of variance (which has removed the variances from PC_1) and is orthogonal to PC_1 (the correlation between PC_1 and PC_2 is zero). With every principal component being independent of each other, the problem of multicollinearity will not occur in PCA (Taffler, 1983). More principal components are constructed on the basis of orthogonal with other principal components using the same approach. In this research, the first five principal components were selected for the final model construction. The total variance explained in each variable group among different time scales are presented in Table 6.6:

Table 6.6 Total Variance Explained in Each Variable Group

	M1	M2	M3	M4	M5
G1	56.91%	58.92%	62.17%	62.95%	57.21%
G12	57.99%	60.82%	63.79%	63.33%	57.31%

For the interpretation of each principal component, the eigenvector plays an important role. The eigenvector are the weights of the variable in the principal component. Therefore, the higher the *'absolute'* value of a variable's eigenvector, the more impact it has on the principal component. Comrey and Lee (1992) argued that an absolute value of above 0.71 (that is, overlapping 50% variance) implies that the variable is extremely useful for explaining the principal component. Table 6.7 summarizes the significant variables with strong explanatory power in each variable group among different time scales.

Table 6.7 Significant Variables in Each Principal Component

G1 (M1)		G12 (M1)	
PC1	V1, V2, V3, V4, V6, V8, V10, V12, V22, V26	PC1	V1, V2, V3, V4, V6, V8, V10, V12, V22, V26

Table 6.7 Significant Variables in Each Principal Component (Continued.)

G1 (M1)		G12 (M1)	
PC2	V36, V37, V38, V39, V41	PC2	V56, V58, V59, V62, V63, V64, V65, V66, V67
PC3	V33	PC3	V36, V37, V38, V39, V41
PC4	V19, V20	PC4	V19, V20, V21
PC5	V5, V9	PC5	V57, V60
G1 (M2)		G12 (M2)	
PC1	V1, V2, V3, V4, V6, V8, V10, V12, V53	PC1	V56, V58, V59, V60, V62, V63, V64, V65, V66, V67
PC2	V36, V37, V38, V39, V41	PC2	V1, V2, V3, V4, V6, V8, V10, V12, V53
PC3	V25, V27	PC3	V36, V37, V38, V39, V41
PC4	V33	PC4	V25, V27
PC5	V5, V9	PC5	V7
G1 (M3)		G12 (M3)	
PC1	V1, V2, V3, V4, V6, V8, V10, V11, V12, V22	PC1	V56, V58, V59, V60, V62, V63, V64, V65, V66, V67
PC2	V36, V37, V38, V39, V41	PC2	V1, V2, V3, V4, V6, V10, V11, V12, V22
PC3	V33, V51	PC3	V36, V37, V38, V39, V41
PC4	V25, V27	PC4	V33, V51
PC5	V52	PC5	V25, V27
G1 (M4)		G12 (M4)	
PC1	V1, V2, V3, V4, V6, V10, V12, V53	PC1	V1, V2, V3, V4, V6, V10, V12, V53
PC2	V36, V37, V38, V39, V41	PC2	V56, V58, V59, V62, V64, V66, V67
PC3	V33, V51	PC3	V36, V37, V38, V39, V41
PC4	V25, V27	PC4	V33, V51
PC5	V52	PC5	V57, V63, V65
G1 (M5)		G12 (M5)	
PC1	V1, V2, V4, V6, V10, V12	PC1	V1, V2, V4, V6, V10, V12
PC2	V36, V37, V38, V39, V41	PC2	V56, V58, V59, V64, V66, V67
PC3	V33, V51	PC3	V36, V37, V38, V39, V41
PC4	V25, V27	PC4	V57, V63, V65
PC5	V46, V47	PC5	V33, V51

Regardless of time periods, results from Table 6.7 can be further rearranged in terms of the frequency of variable appearance. The new results are presented in Table 6.8:

Table 6.8 Rearranged Significant Variables in Each Principal Component

G1		G12	
PC1	Profitability Variables	PC1	Profitability Variables
PC2	Financial Scale Variables	PC2	External Environmental Factors
PC3	Leverage or Brand Strength Variables	PC3	Financial Scale Variables
PC4	Sustainability Variables	PC4	Leverage or Brand Strength Variables
PC5	Productivity or Profitability Variables	PC5	Sustainability, Productivity or Profitability Variables

From Table 6.8, it appears that *Profitability Variables* can explain most variances in the G1 variable group, followed by *Financial Scale Variables, Leverage or Brand Strength Variables, Sustainability Variables and Productivity or Profitability Variables.* With regards to the G12 variable group, *External Environmental Factors* replaced the position of *Financial Scale Variables.* When external influences are added, other variables decreased in significance. Does this imply that macro-environmental factors impact greatly on the performance of default prediction models? This question will be discussed in Chapter Eight.

PCA default prediction models were constructed in terms of the scores of principal components. The component scores can be calculated by multiplying the component score coefficient matrix and the matrix of standardize independent variable values. The scores of the first five principal components will be the primary data for PCA default prediction model development.

6.3.5 Stepwise Regression Approach

As with univariate analysis, logistic regression was selected to carry out the procedure of stepwise variable selection. Hosmer and Lemeshow (2000) argued that *'Employing a stepwise selection procedure can provide a fast and effective means to* *screen a large number of variables, and to fit a number of logistic regression equations simultaneously.'*

There are two kinds of stepwise selection strategies for variable selection: *Forwards Stepwise Approach* and *Backwards Stepwise Approach*. Forwards stepwise approach selects the most statistically significant variable into the model at each step. In contrast, backwards stepwise approach eliminates the statistically most insignificant variables from the model at each step. Therefore, forwards stepwise approach focuses on the key variables selection, whilst backwards stepwise approach concentrates on the unimportant variables elimination. Forward stepwise approach is the main variable selection method here, for the results from the backward stepwise approach tend to overfit the model, especially in context of limited data. The key selected variables in each variable group among different time periods are illustrated in Table 6.9:

Table 6.9 Key Variables in Each Variable Group

Internal Resource Group (G1)	
M1	V2, V6, V10, V15, V20, V33, V49
M2	V5, V12, V29, V33, V42
M3	V6, V16, V29, V33, V42, V53
M4	V16, V33, V37, V41, V42
M5	V4, V8, V33, V36, V42, V46, V52
Internal Resource plus External Factors Group (G12)	
M1	V2, V6, V10, V15, V33, V55, V56, V57
M2	V5, V12, V29, V33, V42
M3	V31, V33, V42, V60, V65
M4	V8, V26, V29, V37, V53, V54, V61, V65
M5	V33, V37, V42, V65

The final key variables in each variable group were selected in terms of the criterion of the variable's appearance frequency. For final model construction, the top five variables with higher appearance frequency in each variable group were selected; these are shown in Table 6.10. This parallels the five principal components used in analysis and hence allows a fair comparison.

Table 6.10 Key Performance Measures

Variable Group	Key Performance Measures
G1	V6: Net Profit Margin V16: Payables Turnover V29: Debt Ratio V33: Total Debt / (Total Debt + Market Capitalization) V42: Operation Cash Flow
G12	V29: Debt Ratio V33: Total Debt / (Total Debt + Market Capitalization) V37: Total Assets V42: Operation Cash Flow V65: Five Years Correlation Coefficient between Government Debt and Total Sales

Adding external influences leads to the replacement of *Net Profit Margin (V6)* and *Payables Turnover (V16)* by *Five Years Correlation Coefficient between Government Debt and Total Sales (V65)* and *Total Assets (V37)*. Following are discussions of the key variables.

6.3.5.1 Net Profit Margin

Net Profit Margin = Net Profits / Total Sales (6.2)

Net profit margin reflects a company's final performance in profitability. Moreover, net profits indicate the rewards to stockholders (dividends) or to the company itself (retained earnings). Therefore, most companies set an appropriate level of net profit margin to be their primary annual objective. However, unlike the operating profit margin, net profit margin does not focus on a company's primary business line and it is possible to give the wrong impression about a company's profitability. For example, if a retailer does not perform well in selling retailing goods but instead, has many non-operating profits, such as capital gains from financial investment, the net profit margin will only reflect the good performance based on results from the financial operations. A high net profit margin in this case does not mean that this company has good profitability. In fact, this company is very risky, since it does not perform well in terms of its primary business line. As a result, if researchers want to explore a company's profitability, they had better take into account different profitability measures to obtain a more objective conclusion.

6.3.5.2 Payables Turnover

Payables Turnover = Cost of Good Sold / Average Accounts Payable (6.3)

Payables turnover measures a company's ability to pay off its trade creditors. A high ratio implies that the company pays its trade creditors very often. In other words, the company cannot keep its cash flow for long and is likely to face higher pressure from the cash flow operation. Payables turnover can be employed to evaluate the relationship between supplier and retailer. If a retailer has good relationship with its suppliers, the payable turnover will be lower. This implies that the retailer has longer buffer time to pay the cost of purchases from its suppliers.

6.3.5.3 Debt Ratio and Total Debt / (Total Debt + Market Capitalization)

Debt Ratio = Total Debts / Total Assets (6.4)

Debt ratio and total debt / (total debt + market capitalization) are used to evaluate a company's leverage situation, especially its ability to face long-term obligations. Hence, the two measures are related to a company's credit assessment directly. One of the differences between these two measures is equity evaluation. For debt ratio (total debts/total assets), the value of equity is evaluated by the accounting value, whilst the equity value is evaluated by market value for another leverage measure.

Another difference between these two measures is the maximum value. For the ratio of total debt / (total debt + market capitalization), the maximum value is one, since the minimum value of the market capitalization is zero. However, for the debt ratio, the maximum value may be greater than one, since the value of total debt may be greater than the value of total assets. In other words, the company would not be able to cover its future obligations even if it sells all its assets. In fact, if a company's debt ratio is greater than one, then it runs into a stock-based insolvency situation (Ross et al., 1999), or, financial distress.

6.3.5.4 Total Assets (log)

As mentioned in Chapter Two, scale measures are more important in the retail industry than in other industries, as one of the important characteristics in the retail industry is low-margin. Large firms usually have certain advantages that small firms do not. For example, large firms have better risk endurance when the economic situation changes. Moreover, large firms also have better financial flexibility, since they can more easily ask for a loan from a financial institution compared with small firms (S&P, 2003). Therefore, financial scale is a significant variable for evaluating a retailer's credit risk.

6.3.5.5 Operating Cash Flow (log)

Sustainability measures a company's ability to service external sources of finance, such as interests. S&P (2003) pointed out that a company's sustainability must be based on cash flow, rather than on earnings in the accounting statements, for earnings include non-cash items that cannot reflect a company's ability to pay back future obligations. As a result, if a company has adequate operating cash flow, then its default risk will be lower.

6.3.5.6 Government Debt / GDP

Government debt / GDP can be regarded as a measure to evaluate a country's leverage situation. It indicates the ability of a country to cover its total debt by using GDP. A number of credit rating agencies use this measure to evaluate a country's sovereign risk (S&P, 2005). In order to assess this measure's impact on each sample company, a five years correlation coefficient between government debt / GDP and total sales is employed in this study.

6.4 Concluding Remarks

This chapter introduced two key issues on the topic of developing a default prediction model: data collection and key variables selection. With regards to data collection, 67 variables in the *'quantifiable measure and available data'* group were selected as the main measures for final model construction, as these measures could be easily calculated and collected from financial database, each sample company's annual report, or other secondary data sources. With the aim of detecting the external environmental influences, these 67 variables were regrouped into two variable groups: *Internal Resources Group (G1)* and *External Factors Group (G2)*. The sample selection criteria for both healthy and distressed firms were also introduced. For example, the selection criterion for distressed firms was based on a legal point of views: any company filing for the bankruptcy code of Chapter 11 or Chapter 7 were deemed to be under financial distress. Finally, a data set of 195 US healthy retailers and 51 US distressed retailers with 67 variables over five time periods from 1994 to 2002 were collected.

To avoid over-fitting with a large number of variables the number of variables was reduced and key variables for final default prediction model construction determined. (Prior to determining the final key variables or principal components, some considerations had been addressed in order to ensure the quality of the final key variables, such as, time-scale consideration, outlier elimination and preliminary univariate analysis.) The most two popular approaches to reducing the number of variables in the financial distress prediction domain are: *Stepwise Regression Approach* and *Principal Component Analysis (PCA)*. These were carried out. The top five principal components and the top five important variables in each variable group were consequently selected for final default prediction model construction.

The next chapter will continue to discuss the key issues regarding the development of a default prediction model. The focus will be on the introduction of credit scoring techniques. Moreover, the cross-validation process will also be presented with the purpose of avoiding the overfitting problem for the final models.

Chapter SEVEN

Model Construction: Modelling Techniques and Cross-Validation Process

7.1 Introduction

This chapter will continue to discuss the key issues relative to the development of a default prediction model, with a special focus on credit scoring techniques. In addition, this chapter will discuss the cross-validation process that was carried out to avoid overfitting. The cross-validation procedure will be introduced in Section 7.3. In Section 7.4, the software for model construction will be introduced. The final section summarises the key issues in this chapter.

7.2 Credit Scoring Techniques

In this research, five credit scoring methodologies: *Naïve Bayes*, *Logistic Regression*, *Recursive Partitioning*, *Artificial Neural Network*, and *Sequential Minimal Optimization* were employed for final default prediction models construction. The following sections will introduce these five techniques.

7.2.1 Naïve Bayes

The default prediction problem can be regarded as a problem of evaluating the probability of financial distress conditional upon the value of a specific financial ratio (Beaver, 1966). *Naïve Bayes* approach provides a straightforward method to deal with a classification problem. Let H be the healthy samples and D be the distressed samples. The conditional probability of a financial distress firm or a healthy firm in terms of a specific value of financial ratio X can be expressed as: $P(D \mid X)$ and $P(H \mid X)$.

In order to work out these two conditional probabilities, the prior probabilities $P(D)$ and $P(H)$ should be calculated in advance. For example, if a data set contains 200 healthy firms and 50 distressed firms, the prior probabilities can be calculated as: $P(D) = 50 / 250 = 0.2$ and $P(H) = 200 / 250 = 0.8$ ($P(D)$ plus $P(H)$ is equal to one). The next step is to calculate the conditional probabilities based on a specific financial ratio X for both healthy and distressed samples: $P(X \mid H)$ and $P(X \mid D)$. For example, assume X is the current ratio with a value of 2.5. If there are 50 healthy firms and 10 distressed firms with a current ratio of 2.5, then $P(X \mid H)$ and $P(X \mid D)$ can be calculated as follows:

$$P(X = 2.5 \mid H) = \frac{P(X = 2.5 \cap H)}{P(H)} = \frac{50 / 250}{0.8} = 0.25 \tag{7.1}$$

$$P(X = 2.5 \mid D) = \frac{P(X = 2.5 \cap D)}{P(D)} = \frac{10 / 250}{0.2} = 0.2 \tag{7.2}$$

Where the $P(X = 2.5 \cap H)$ and the $P(X = 2.5 \cap D)$ are the joint probabilities and sum of these two joint probabilities is the marginal probability $P(X)$—the probability of having a current ratio of 2.5 could occur, for: $P(X) = (50 / 250) + (10 / 250) = 0.24$. Drawing on above, the $P(D \mid X)$ and $P(H \mid X)$ can be calculated as follows:

$$P(H \mid X = 2.5) = \frac{P(X = 2.5 \cap H)}{P(X)} = \frac{50 / 250}{0.24} = \frac{5}{6} \tag{7.3}$$

$$P(D \mid X = 2.5) = \frac{P(X = 2.5 \cap D)}{P(X)} = \frac{10 / 250}{0.24} = \frac{1}{6} \tag{7.4}$$

Obviously, the sum of $P(D \mid X)$ and $P(H \mid X)$ is equal to one. The calculations above can be simplified by using the rule of *Naïve Bayesian* that copes with the classification problem.

$$P(H \mid X) = \frac{P(H)P(X \mid H)}{P(X)} = \frac{0.8 \times 0.25}{0.24} = \frac{5}{6} \tag{7.5}$$

$$P(D \mid X) = \frac{P(D)P(X \mid D)}{P(X)} = \frac{0.2 \times 0.2}{0.24} = \frac{1}{6} \tag{7.6}$$

7.2.2 Logistic Regression

Unlike the traditional linear regression, the dependent variable in the logistic regression is dichotomous. In this research, the dependent variable can be expressed dichotomously: *'1'* for healthy firms and *'0'* for distressed firms. Why logistic regression should be employed to deal with the problem of a regression with a dichotomous dependent variable is explained below.

In any regression, the dependent variable can be expressed as *E (Y | X)*, which means the expected value of *Y* given the value *X*. Therefore, a simple linear regression model can be presented as:

$$E(Y \mid X) = \beta_0 + \beta_1 X \tag{7.7}$$

As the value of *X* ranges between $-\infty$ and $+\infty$, it is possible that the *E (Y | X)* takes any value. However, as the dependent variable is dichotomous in the default prediction domain, the value of *E (Y | X)* should be between 0 and 1. In addition, the relationship between the *E (Y | X)* and *X* may not be linear. Hence, the traditional *Normal Distribution* assumption for the linear regression becomes inappropriate for coping with dichotomous dependent variables. In contrast, the logistic distribution can be used to solve this problem.

The logistic distribution is an *S-shape* distribution. The change in the *E(Y | X)* in terms of a unit change in X becomes increasingly smaller when the *E(Y | X)* gets closer to zero or one (Hosmer and Lemeshow, 2000). Therefore, the range of *E (Y | X)* is between 0 and 1. The logistic distribution can be expressed for a simple variable x_i as follow:

$$l(x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_i)}} = \frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}} = \frac{e^{z_i}}{1+e^{z_i}} \tag{7.8}$$

$$1-l(x) = 1-\frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}} = \frac{1}{1+e^{\beta_0+\beta_1 x_i}} = \frac{1}{1+e^{z_i}} \tag{7.9}$$

where $Z_i = \beta_0 + \beta_1 x_i$. Since Z_i is between $-\infty$ and $+\infty$, $l(x)$ will be between 0 and 1 ($Z_i \rightarrow -\infty$, $l(x) \rightarrow 0$; $Z_i \rightarrow +\infty$, $l(x) \rightarrow 1$). Moreover, it is obvious that the relationship between $l(x)$ and x_i is not linear. Therefore, logistic distribution is more appropriate for describing a regression with a dichotomous dependent variable. By a logit transformation on odd ratio function, the logistic model can be linearized and be used to solve classification problems.

$$\frac{l(x)}{1 - l(x)} = \frac{e^z}{1 + e^z} \times \frac{1 + e^z}{1} = e^z = e^{\beta_0 + \beta_1 x_i} \tag{7.10}$$

$$g(x) = \ln\left[\frac{l(x)}{1 - l(x)}\right] = \beta_0 + \beta_1 x_i \tag{7.11}$$

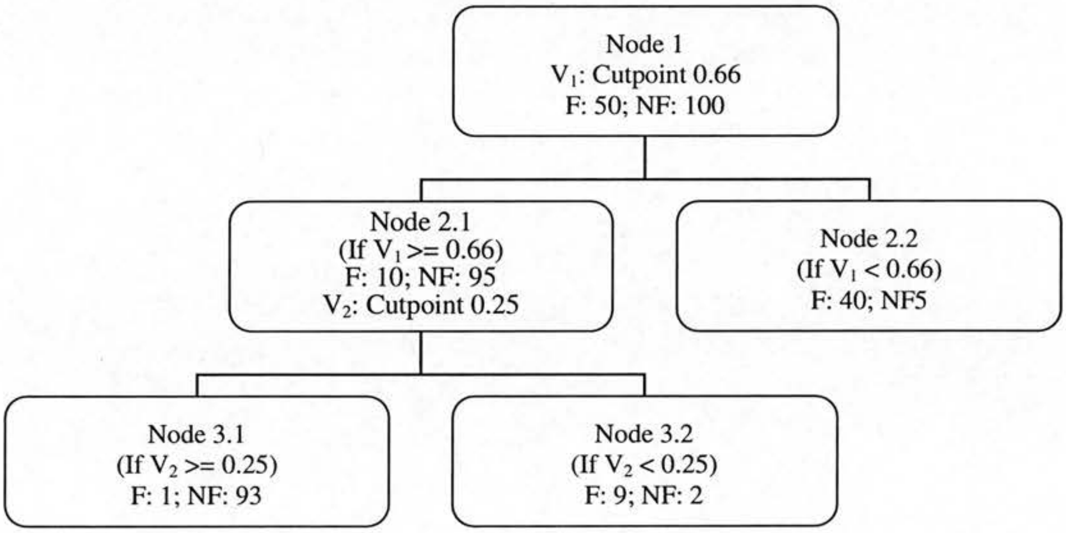
From Function 7.11, it is clear that although the relationship between $g(x)$ and x_i is linear, the relationship between $l(x)$ and x_i is still nonlinear. Moreover, although the value of $l(s)$ is between 0 and 1, the value of $g(x)$ can take any value and hence $g(x)$ will not be bounded. By conducting the *Maximum Likelihood* method, the β_0 and β_1 can be estimated. Consequently, logistic regression can appropriately classify healthy and distressed firms.

7.2.3 Recursive Partitioning

RPA is a nonparametric technique and hence, the dependent and independent variables do not require any distribution assumptions. RPA can be viewed as a stepwise procedure. The first step is to select an independent variable to be the best discriminator and to decide a cutpoint value in terms of the lowest expected misclassification cost. For example, assume the number of failed firms *(F)* and nonfailed firms *(NF)* are 50 and 100 respectively and V_1 is selected to be the best discriminator with a cutpoint 0.66 (see Figure 7.1, node 1). Based on the cutpoint, the second step is to divide both failed and nonfailed firms into the node 2.1 and node 2.2. Because there are still a number of mixing firms in node 2.1, further partitioning is required. The third step is to select another (or the same) discriminator, V_2, to partition the failed and nonfailed firms in node 2.1 into node 3.1, and to partition node 3.2 using a specific cutpoint value (the value 0.25 is assumed). The same process can be continued, if further splitting is necessary.

As mentioned in Chapter Two, Thomas et al. (2002) mentioned two reasons to stop the partitioning process. The first reason is, the number of samples in a node is too small, making it unnecessary to partition them. The second reason is, if the classification results between the old node and new nodes do not have significant differences, then it also not necessary to split the old node. Assuming the partitioning process ends at this step, node 2.2, node 3.1 and node 3.2 can be identified as the terminal nodes.

Figure 7.1 An Example of Recursive Partitioning



The performance of classification can be described as follows: in node 2.2, 40 distressed firms are classified correctly and five healthy firms are misclassified; in node 3.1, 93 healthy firms are classified correctly and one failed firm is misclassified; in node 3.2, nine failed firms are classified correctly and two healthy firms are misclassified. The classification matrix is presented in Table 7.1. The overall classification accuracy rate is (49+93) / 150 = 94.67%.

Table 7.1 Classification Matrix

	Failed Firms (F)	Nonfailed firms)NF)	Total
Failed Firms (F)	49 (40+9)	1	50
Nonfailed firms (NF)	7 (2+5)	93	100
Total	56	94	150

7.2.4 Artificial Neural Network

The most popular artificial neural network algorithm in the financial distress prediction domain is the *Multilayer Perceptron (MLP).* A MLP has three primary components: input layer, hidden layer and output layer and an example of a three-layered MLP is presented in Figure 7.2.

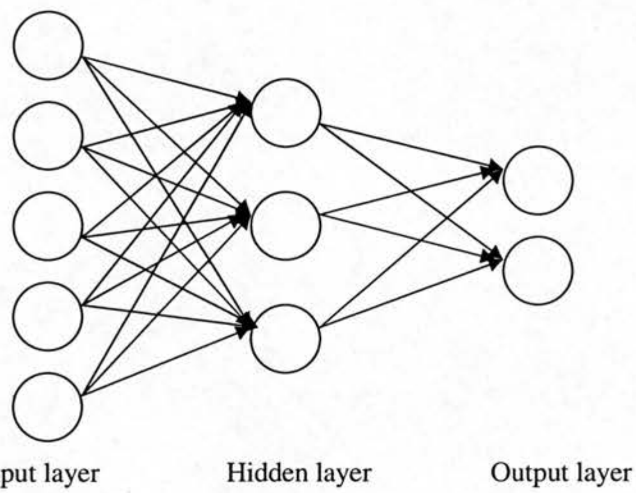
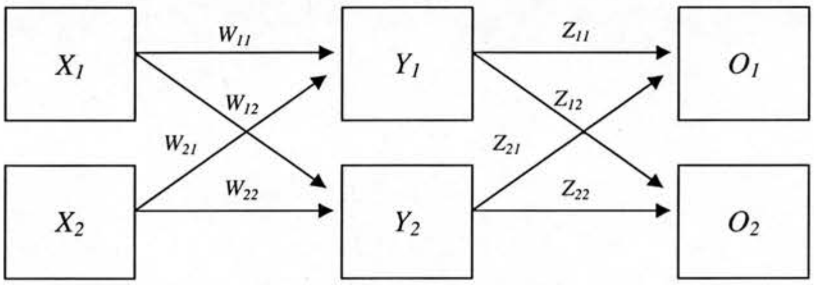


Figure 7.2 Three Layers Multilayer Perceptron

In Figure 7.2, the input layer is responsible for receiving information from the outside environment and transferring it to the hidden layer. In the hidden layer, a neuron will assign a series of weights to the inputs, cope with the information via a training process, and then forward the results with weights to the output layer. The training process can be viewed as the weight determination procedure and the most frequently used algorithm for the training process is the *Back Propagation Algorithm (BPA).* Thomas et al. (2002) pointed out that BPA has two phases: *forward pass* and *backward pass*. The BPA training process can be illustrated by the following example.

Figure 7.3 is a MLP with three layers and each layer includes two neurons. X_1 and X_2 are the inputs of first neuron and second neuron in the input layer; Y_1 and Y_2 are the values of first neuron and second neuron in the hidden layer; O_1 and O_2 are the outputs of the first neuron and second neuron in the output layer. W indicates the weight between the input layer and the hidden layer; Z indicates the weight between the hidden layer and output layer. Therefore, W_{11} represents the weight between the first neuron in the input layer and the first neuron in the hidden later; Z_{11} represents the weight between the first neuron in the hidden layer and the first neuron in the output layer, and so on.

Figure 7.3 An Example for Back Propagation Algorithm



At first, all weights are set to be equal. Y_1 and Y_2 in the hidden layer can be calculated by the following equation:

$$Y_1 = W_{11} X_1 + W_{21} X_2 = \sum_{i=1}^{2} W_{i1} X_i \tag{7.12}$$

$$Y_2 = W_{12} X_1 + W_{22} X_2 = \sum_{i=1}^{2} W_{i2} X_i \tag{7.13}$$

Functions (7.12) and (7.13) can be generalized as:

$$Y_k = \sum_{i=1}^{k} W_{ik} X_i \tag{7.14}$$

In Function (7.14), k is the number of the neuron. The same can be concluded in the output layer; O_1 and O_2 can be obtained by using the formula (7.15):

$$O_k = \sum_{i=1}^{k} Z_{ik} Y_i = f(Y_k) \tag{7.15}$$

Up to now, the procedure is called *forward pass. Backward pass* begins by calculating the difference between the expected output value O_k and the observed output value A_k (called *error*) in the output layer. The error (e_k) can be expressed in Function (7.16):

$$e_k = A_k - O_k \tag{7.16}$$

The main purpose of BPA is to distribute the error back to the network and to adjust the weight to reduce the average error. The process is repeated for all cases, called an *epoch*. After several epochs training, the average learning error will reduce to a minimum level and the training process will end. The average learning error is presented as follow:

$$E(e_k) = \frac{1}{2}\sum_{k=1}^{2} e_k^2 \tag{7.17}$$

As the aim is to adjust the weights to reduce the average learning error, the partial derivative of $E(e_k)$ with respect to weight W_{ik} is carried out by using the chain rule as presented in Function (7.18): (Thomas et al., 2002: 74).

$$\frac{\partial E(e_k)}{\partial W_{ik}} = \frac{\partial E(e_k)}{\partial e_k} \times \frac{\partial e_k}{\partial O_k} \times \frac{\partial O_k}{\partial Y_k} \times \frac{\partial Y_k}{\partial W_{ik}} \tag{7.18}$$

$$\frac{\partial E(e_k)}{\partial e_k} = e_k \tag{7.19}$$

$$\frac{\partial e_k}{\partial O_k} = -1 \tag{7.20}$$

$$\frac{\partial O_k}{\partial Y_k} = f'(Y_k) \quad (7.21)$$

$$\frac{\partial Y_k}{\partial W_{ik}} = X_i \quad (7.22)$$

Hence,

$$\frac{\partial E(e_k)}{\partial W_{ik}} = -e_k \times f'(Y_k) \times X_i \quad (7.23)$$

Based on the Widrow-Holf learning rule (Gluck and Myers, 2001:53), the change in weight can be presented in Function (7.24):

$$\Delta W_k = -\eta \times \frac{\partial E(e_k)}{\partial W_{ik}} = \eta \delta_k X_i \quad (7.24)$$

where $\delta_k = e_k \times f'(Y_k)$ (7.25)

The η is the learning rate. Gluck and Myers (2001) pointed out that the learning rate is a fixed small number, which determines the amount of change in weight based on a single trial. Thomas et al. (2002) argued that smaller training rates can improve the training accuracy, but increase the training time. Therefore, the learning rate can be viewed as a trade-off between prediction performance and training cost.

Normally, most implementations consider adding a *momentum term* in the Function (7.24) with the aim of improving the speed of training. Under this situation, the change of weight will be affected not only by current error, but also by previous error. The new function added momentum is presented in the equation (7.26):

$$\Delta W_k = \eta \delta_k X_i + \Delta W_K^* \quad (7.26)$$

where

ΔW_K^* is the amount of weight change in the previous trial

7.2.5 Sequential Minimal Optimization

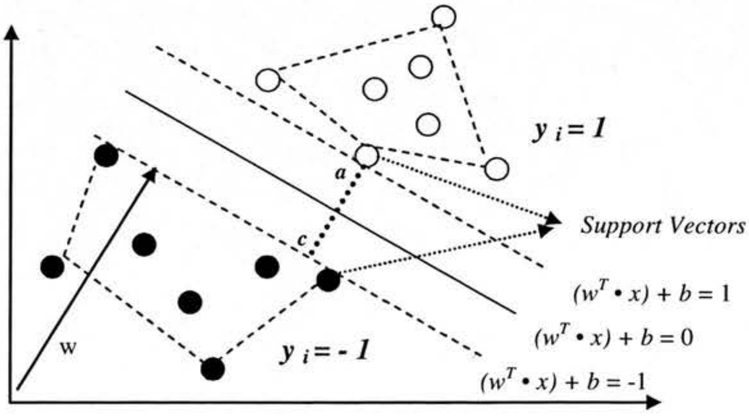
Sequential Minimal Optimization (SMO) was developed by Platt (1999) and it is a special form of *Support Vector Machine (SVM)*. Platt argued that a large number of *Quadratic Programming (QP)* in SVM training is time consuming and too complex to implement in the real world. SMO can be used to improve the SVM training time.

SVM was applied to default prediction research in the late 1990s (Fan and Palaniswami, 2000). The primary advantage of SVM is that it is able to classify healthy and distressed firms based on some complex data patterns by generating a highly nonlinear separating surface. Assuming the research intends to use n performance measures to classify l firms into healthy firms and distressed firms. The performance measures of the i^{th} firm can be presented as an input vector $x_i = (x_1, x_2 \ldots x_n)$. The dependent variable (or the target label) is expressed in binary form; $y_i = 1$ indicates a healthy firm and $y_i = -1$ means a distressed firm. The whole dataset D can be presented as follow:

$$D = \{(x_1, y_1), (x_2 y_2), \ldots, (x_l, y_l)\} \tag{7.27}$$

The support vector classifier can be displayed as the line $(w^T \bullet x) + b = 0$ in Figure 7.4 (Hearst, 1998). where the w^T is the weight vector and b is the threshold.

Figure 7.4 Separating Hyperplane



In Figure 7.4, all cases can be completely classified by the line: $(w^T \bullet x) + b = 0$. Moreover, the cases closest to the hyperplane are defined as *Support Vectors*. The basic decision functions are:

$$\begin{cases} (w^T \bullet x) + b \geq 1, & if \;\; y_i = 1 \\ (w^T \bullet x) + b \leq -1, & if \;\; y_i = -1 \end{cases} \tag{7.28}$$

where i = 1, 2, … ,l.

A healthy company will be in the area $(w^T \bullet x) + b \geq 1$, whilst a distressed company will be in the area $(w^T \bullet x) + b \leq -1$. Function 7.28 can be re-written and it is equivalent to:

$$y_i((w^T \bullet x) + b) \geq 1 \tag{7.29}$$

The primary objective of SVM is to maximise the margin between healthy firms and distressed firms, that is, to maximise the distance between point *a* and point *c* in Figure 7.4. Based on the concept of Euclidean distance, the distance between *a* and *c* can be presented as Function (7.30):

$$\overline{ac} = \frac{2}{\sqrt{w^T \bullet w}} = \frac{2}{\|w\|} \tag{7.30}$$

Maximizing Function (7.30) is the same as minimizing the function of $\frac{1}{2}\|w\|^2$. In addition, a slack variable can be included in order to deal with non-separable situations. Finally, the optimal problem can be illustrated in the Function (7.31):

$$\min \;\; \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l} \xi_i \tag{7.31}$$

$$\text{Subject to} \begin{cases} y_i((w^T \bullet x) + b) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases} \tag{7.32}$$

The ξ_i is the slack variable, which allows margin misclassifications (Fan and Palaniswami, 2000). C is the tuning hyperparameter, which controls the trade-off between classification ability and training errors. This optimal problem can be solved by building the Lagrangian model based on the KKT conditions. After constructing the Lagrangian model, the optimal problem can be transformed into a dual problem, which is identical to maximizing:

$$\max \quad Q(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j (x_i^T, x_j) \tag{7.33}$$

$$\text{Subject to} \begin{cases} 0 \leq \alpha_i \leq C \\ \sum_{i=1}^{i} \alpha_i y_i = 0 \end{cases} \tag{7.34}$$

Where the term of (x_i^T, x_j) is the linear dot product. However, in the real world, most problems are not linearly separable. With the aim of classifying healthy and distressed firms based on the highly nonlinear separating surface, the *Kernel* functions were employed to replace the linear dot product. Drawing on this insight, the Function (7.33) can be revised as follow:

$$\max \quad Q(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \tag{7.35}$$

$$\text{Subject to} \begin{cases} 0 \leq \alpha_i \leq C \\ \sum_{i=1}^{l} \alpha_i y_i = 0 \end{cases} \tag{7.36}$$

The non-linear decision function can be presented in the equation (7.37):

$$u = \sum_{i=1}^{l} y_i \alpha_i k(x, x_i) + b \tag{7.37}$$

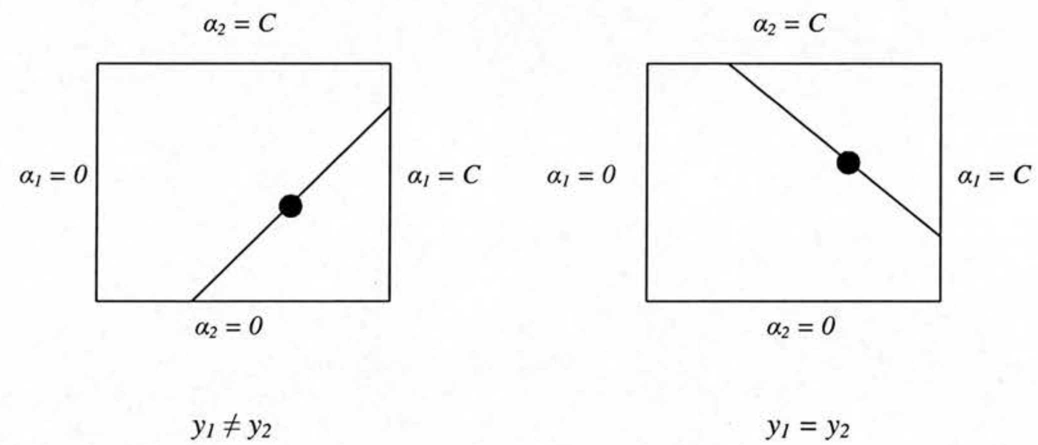
$$f(x) = \operatorname{sgn}(u), \tag{7.38}$$

where $f(x)$ is the class of the predicted label for the x

In Function (7.35), the primary quadratic programming (QP) problem is to solve the Lagrange multipliers: α_i. However, it is clear that a large number of calculations are required to compute α_i, since each training sample x corresponds to a Lagrange multiplier. Platt (1999) pointed out that SMO is designed to cope with this problem by only using two Lagrange multipliers (α_1 and α_2) at each training step. As a result, SMO can be viewed as a decomposition method with the aim of decomposing the overall QP problems into fixed-size sub-problems.

Function (7.36) presents the two constraints for the Function (7.35). With the first inequality constraint $0 \leq \alpha \leq C$, the two Lagrange multipliers α_1 and α_2 lie in the two boxes shown in Figure 7.5 (Platt, 1999:189):

Figure 7.5 Two Lagrange Multipliers of Optimizations



The second equality constraint based on the two Lagrange multipliers α_1 and α_2 can be expressed as follows:

$$\sum_{i=1}^{l} \alpha_i y_i = 0 \rightarrow \alpha_1 y_1 + \alpha_2 y_2 + \sum_{i=3}^{l} \alpha_i y_i = 0$$

$$\rightarrow \alpha_1 y_1 + \alpha_2 y_2 = \delta \qquad (7.39)$$

$$\text{where } \delta = -\sum_{i=3}^{l} \alpha_i y_i \qquad (7.40)$$

Since $y_i \in \{1,-1\}$, hence: (Platt, 1999:189)

$$\text{If } y_1 \neq y_2, \qquad \alpha_1 - \alpha_2 = \delta \tag{7.41}$$

$$\text{If } y_1 = y_2, \qquad \alpha_1 + \alpha_2 = \delta \tag{7.42}$$

Let $s = y_1 y_2$, then the Functions (7.41) and (7.42) can be re-written as:

$$\alpha_1 + s\alpha_2 = \delta \rightarrow \alpha_1 = \delta - s\alpha_2 \tag{7.43}$$

Let $v_i = \sum_{j=3}^{l} y_i \alpha_i k_{ij}$, (where $k_{ij} = K(x_i, x_j)$) and put $\alpha_1 = \delta - s\ \alpha_2$ and α_2 into Function (7.35), we get: (Platt, 1999:204)

$$Q(\alpha_1 = \delta - s\alpha_2, \alpha_2) = \delta - s\alpha_2 + \alpha_2 - \frac{1}{2}k_{11}\delta^2 + k_{11}s\delta\alpha_2 - \frac{1}{2}k_{11}\alpha_2^2$$
$$-\frac{1}{2}k_{22}\alpha_2^2 - sk_{12}\delta\alpha_2 + k_{12}\alpha_2^2 - y_1v_1\delta + y_1v_1s\alpha_2 - y_2v_2\alpha_2 + Con. \tag{7.44}$$

By maximizing the Function (7.44), the new α_2 can be obtained by Function (7.45):

$$\alpha_2^{new} = \alpha_2 + \frac{y_2(E_1 - E_2)}{k_{11} + k_{22} - 2k_{12}} \tag{7.45}$$

where $E = u_i - y_i$, which is the error between the output of SVM (u_i) and the class label of x_i (y_i). α_1 can be calculated as follows:

$$\alpha_1^{new} + s\alpha_2^{new} = \alpha_1 + s\alpha_2 \rightarrow \alpha_1^{new} = \alpha_1 + s(\alpha_2 - \alpha_2^{new}) \tag{7.46}$$

Drawing on above, it can be concluded that as SMO only solves two Lagrange multipliers at each training step, it does not require the numerical quadratic programming optimization process or any extra matrix storage. Hence, the training time can be improved. Thus far, this research has introduced the five credit scoring techniques for final model construction. The next section looks at a cross-validation process carried out with the objective of avoiding potential overfitting prior to final model construction.

7.3 Cross-Validation Approach

Three cross-validation methods will be discussed in this section: the test set method, the leave-one-out method and the 10-folders method. The test set method is the traditional cross-validation approach. It begins by randomly selecting 30% of whole dataset to be the test set, and another 70% of the data to be the training set. After classification in terms of the training set data, the test set is employed to estimate the classification error. If the estimated classification error is high, then the results from the training set are potentially overfitting. The main advantage of the test set method is simplicity and low cost. However, if the sample size is small, the estimated classification error tends to have high variance. In addition, test set method will also waste 30% of the data to perform the cross-validation procedure.

The leave one-out-method is the same as the *Lachenbruch* method (Lachenbruch, 1975). Assume the number of total observations is n. One of the total observations will be removed temporarily and the remaining $n-1$ observations will be trained in each training process. After n times training, the average classification error is calculated to estimate the model overall performance. Lachenbruch (1975) pointed out that the leave-one-out method can provide an almost unbiased estimate of model classification ability and will not waste any data (Moore, 2001). However, leave-one-out method is only appropriate if the sample size is very small, since it is time-consuming and costly.

Moore (2001) introduced the 10-folders cross-validation method in order to address the advantages of both test set approach and leave-one-out approach. The basic idea is that the original data set is first randomly divided into to 10 folders, followed by training on each folder. The average performance of these 10 training folders is employed to estimate the overall model classification performance. Therefore, the 10-folders cross-validation approach only wasted 10% of total data and the training cost was much lower than the leave-one-out method. This research will adapt the 10-folders approach to be the primary cross-validation technique.

7.4 Software for Model Construction

The primary software used to create the default prediction model was the Java Machine Learning Software: *Weka*[1]. Weka was designed by the University of Waikato and includes several main functions in the data-mining domain, such as, *classification*, *association*, and *clustering*. In addition, it provides several cross-validation techniques, such as the test set method and the *k*-folders method. Weka software is a free open source, which is available under the General Public License.

7.5 Concluding Remarks

This research employed five credit scoring techniques: *Naïve Bayes*, *Logistic Regression*, *Recursive Partitioning*, *Artificial Neural Network*, and *Sequential Minimal Optimization* for final default prediction models construction. In particular, the *Sequential Minimal Optimization* technique is a very new approach for forecasting corporate financial distress and thus far, not many studies have applied it to corporate default prediction.

Prior to the final model construction, this research also carried out the cross-validation process in order to avoid overfitting problems. This research chose the 10-folders cross-validation method, as the 10-folders cross-validation approach only wasted 10% of total data and the training cost was much lower than the leave-one-out method.

Given that this research considered two variable selection methods (*Forward Stepwise Approach* and *Principal Component Analysis*), two different variable groups (*G1* and *G12*), five different time periods (from *M1* to *M5*), and five credit scoring modelling techniques, a total number of 100 models were constructed.

---

1 Downloadable from website: http://www.cs.waikato.ac.nz/~ml/weka/

From the next chapter onwards, the researcher will focus on the evaluation of the default prediction model's performance. As mentioned in Chapter One, the utility of default prediction model will be assessed in terms of two criteria: model prediction power and practical applicability. The next chapter will concentrate on the evaluation of the model's prediction utility. In addition, the approaches for evaluating prediction performance will also be introduced.

Part Five

Model Utility Evaluation

- Chapter Eight: ***Model Prediction Performance Evaluation***
- Chapter Nine: ***Model Practical Applicability Evaluation: Comparison with Moody's Rating***
- Chapter Ten: ***Model Practical Applicability Evaluation: International Applicability***
- Chapter Eleven: ***Generic Global Model Development and Performance Evaluation***

Chapter Eight will evaluate the default prediction models' performance by using predictive accuracy rates and AUROC. An introduction of the approaches for model performance measurement will also be presented. Furthermore, some key issues, such as detection of external influences, evaluation of types of error and the exploration of time series effects will also be addressed. Chapters Nine and Ten follow with assessments of the practical applicability of default prediction models. In *Chapter Nine*, the assessment is carried out by comparing the PhD research models with Moody's credit ratings. Data collection for Moody's ranking data and the techniques for comparison are also introduced. In *Chapter Ten,* the assessment of practical applicability consists of applying the model to other market datasets in different time periods. A new US dataset, a European dataset and a Japanese dataset are used. An international comparison analysis of the models' prediction performance is then carried out. *Chapter Eleven* will focus on the development of a composite model based on combining data from US, European and Japanese markets. The prediction ability and practical applicability of the composite model will also be evaluated. Moreover, a comparative analysis between the composite model and the original USA model will also be conducted.

Chapter EIGHT

Model Prediction Performance Evaluation

8.1 Introduction

In Chapters Six and Seven, the focus was on the development of default prediction models. Given that this research took into account two variable selection methods (*Principal Component Analysis* and *Forward Stepwise Approach*), two different variable groups (*G1* and *G12*), five different time periods (from *M1* to *M5*), and five credit scoring techniques, a total number of 100 models were developed.

As mentioned in Chapter One, the research objective is not only to develop default prediction models, but also, to construct *effective* ones. Therefore, model utility evaluation was carried out. A model's performance can be assessed through two criteria: model prediction power and practical applicability. In this chapter, the evaluation of the model prediction power will be the primary focal point.

This chapter begins by introducing the approaches for evaluating model prediction performance. It then progresses to a discussion on the model's prediction performance using two variable selection methods: Principal Component Analysis (PCA) and Forward Stepwise Approach. Key issues, such as those related to time scale, types of error and external influences is also discussed. In addition, a comparative analysis between the PCA and Forward Stepwise approaches will be presented. The final section summarises the key findings of this chapter.

8.2 Approaches for Model Utility Assessment

Performance evaluation of the prediction model was carried out using two approaches, known as, the *Classification Accuracy Rate* (Hand, 1997: 119) and the

Area under the Receiver Operating Characteristics Curve (AUROC) (Thomas et al., 2002: 115).

8.2.1 Classification Accuracy Rate

Classification accuracy rate is a straightforward method employed widely in previous studies on default prediction model evaluation. This method was applied to all five different methodologies in this research and is illustrated by a confusion matrix in Table 8.1:

Table 8.1 Confusion Matrix

		Observed Value		
Predicted Value		$y = 0$	$y = 1$	
	$\hat{y} = 0$	A	B	A+B
	$\hat{y} = 1$	C	D	C+D
		A+C	B+D	A+B+C+D

The overall accuracy rate can be calculated as follow:

$$AR = (A + D)/(A + B + C + D) \tag{8.1}$$

where

0 means distressed sample; *1* means healthy sample

$A + B$ is the total number of distressed firms

$C + D$ is the total number of healthy firms

As mentioned in Chapter Two, it is also interesting to explore the issues of different types of error. The overall accuracy rate is defined as the joint minimization of Type I and Type II misclassification errors. Type I error is defined as the error to classify a distressed firm as a healthy firm, while Type II error is defined as the error to predict a healthy firm as a distressed firm. In other words, the Type I error indicates the percentage of the misclassified distressed firms over total distressed firms, while the Type II error indicates the percentage of the misclassified healthy firms over total healthy firms. As a result, Type I error (T_1) and Type II error (T_2) can be defined in the Functions (8.2) and (8.3):

$$T_1 = B /(A + B) \tag{8.2}$$

$$T_2 = C /(C + D) \tag{8.3}$$

The importance between Type I and Type II errors varies depending on the users of the default prediction model. Therefore, the presentation of different types of error will provide valuable information to different stakeholders for rationalizing the decision-making process.

Of importance is the cutpoint determination. As each sample company will be attributed a credit score after modelling process, all 246 sample companies can be ranked in terms of their credit scores. Moreover, since the number of distressed companies is 51, the cutpoint to distinguish between healthy and distressed firms can be determined by basing it upon the credit score value of the 196th credit rank. Based on this approach, the predicted margins will be equal to the observed margins, that is, $A + C = A + B$ or $B + D = C + D$. The alternative approach is to let the software generate its own cutpoint. Unlike the former method, the predicted margins may not be equal to the observed margins. Tables 8.2 and 8.3 are the examples of accuracy rates of Logistic Regression, Neural Network and SMO stepwise regression models based on these two cutpoint determination approaches in terms of the time period one year before financial distress (M1):

Table 8.2 Accuracy Rates based on Equality of Margins Approach

Methodology (G1 Model)	Accuracy rate	Methodology (G12 Model)	Accuracy rate
Logistic Regression Cutpoint: 0.3016	91.87%	Logistic Regression Cutpoint: 0.3873	92.68%
Neural Network Cutpoint: 11.3685	75.61%	Neural Network Cutpoint: 99.4968	81.30%
SMO Cutpoint: -3.8890	87.80%	SMO Cutpoint: -2.717	91.87%

Table 8.3 Accuracy Rates based on Software Generated Cutpoint Approach

Methodology (G1 Model)	Accuracy rate	Methodology (G12 Model)	Accuracy rate
Logistic Regression	89.84%	Logistic Regression	91.87%
Neural Network	93.09%	Neural Network	90.24%
SMO	89.84%	SMO	90.24%

From Tables 8.2 and 8.3, it is obvious that despite the variable groups (G1 or G12), the accuracy rate based on the Equality of Margins approach is similar to the accuracy rate based on the Software Generated Cutpoint approach in terms of the Logistic Regression and SMO techniques. However, the same cannot be concluded for the Neural Network model. The accuracy rate of Neural Network model based on the Software Generated Cutpoint approach is higher than the accuracy rate based on the Equality of Margins approach, and hence the results from the Software Generated Cutpoint approach display better performance. Not wishing to disadvantage a particular method it was decided to subsequently use Software Generated Cutpoint approach which may slightly bias results in favour of Neural Networks.

8.2.2 Area under the Receiver Operating Characteristics Curve (AUROC)

Another approach to evaluate the utility of a default prediction model is the AUROC value. The Receiver Operating Characteristics Curve (ROC curve) is used to explore the relationship between the sensitivity and 1-specificity through a variety of different cutpoints (Thomas et al., 2002: 115).

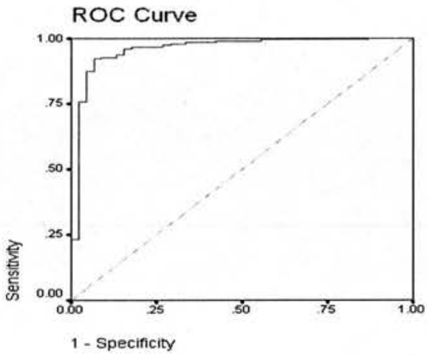
Sensitivity (*Se*) is also called *'True Positive Rate'* and is the probability of predicting as a good company as healthy. Specificity (*Sp*), also called *'True Negative Rate'*, is the probability of a company to be predicted as a distressed company, when this company is truly distressed. The sensitivity and specificity can be calculated based on Table 8.1 as follows:

$$Se = D/(B + D) \quad (8.4)$$

$$Sp = A/(A + C) \quad (8.5)$$

An example of the ROC curve is shown in Figure 8.1.

Figure 8.1 ROC Curve



The area under the ROC curve (AUROC) is the area between the ROC curve and diagonal line and hence the value of AUROC is between 0.5 and 1. The diagonal line of ROC curve reflects the feature of a test with no discriminating power (Hand, 1997:132). In fact, different cut points should reflect different sensitivity and specificity values, since the classification rule is different. Therefore, the further the ROC curve is from the diagonal line, the better the model performance (Thomas et al., 2002: 115). In this research, AUROC was only applied to the Naïve Bayes, Logistic Regression and Neural Network models. SMO does provide a credit score and hence it can produce AUROC results. Unfortunately, the current software implementation does not automatically produce the appropriate AUROC graph and results. Whilst it would be possible to calculate the points on curve, one would still need to measure the AUROC area. As a result and because of the length calculation, it is considered inappropriate to produce the AUROC value. The general rules of the AUROC according to Hosmer and Lemeshow (2000) are presented in Table 8.4.

Table 8.4 General Rules of AUROC

General Rule	Meaning
If AUROC = 0.5	No discrimination
If $0.7 \leq$ AUROC < 0.8	Acceptable discrimination
If $0.8 \leq$ AUROC < 0.9	Excellent discrimination
If AUROC ≥ 0.9	Outstanding discrimination

Source: Hosmer and Lemeshow (2000) Applied Logistic Regression. p. 162

8.3 Prediction Utility Assessment for PCA Models

8.3.1 Classification Accuracy Rate Analysis

The results of the classification accuracy rates based on two variable groups (G1 and G12) among five different time scales are presented in Tables 8.5 and 8.6:

Table 8.5 Classification Accuracy Rates of the PCA Models (G1)

Methodology	Performance	M1	M2	M3	M4	M5	Average
Naïve Bayes	Type I error	64.71%	45.10%	76.47%	54.90%	27.45%	***53.73%***
	Type II error	4.10%	11.28%	9.23%	13.85%	36.92%	***15.08%***
	Overall	83.33%	81.71%	76.83%	77.64%	65.04%	***76.91%***
Logistic Regression	Type I error	37.25%	54.90%	70.59%	60.78%	70.59%	***58.82%***
	Type II error	5.13%	3.59%	6.15%	5.13%	7.18%	***5.44%***
	Overall	88.21%	85.77%	80.49%	83.33%	79.67%	***83.49%***
Neural Network	Type I error	41.18%	50.98%	68.63%	58.82%	70.59%	***58.04%***
	Type II error	8.21%	6.67%	7.69%	3.59%	7.18%	***6.67%***
	Overall	84.96%	84.15%	79.67%	84.96%	79.67%	***82.68%***
SMO	Type I error	90.20%	98.04%	100.00%	100.00%	100.00%	***97.65%***
	Type II error	0.51%	0.00%	0.00%	0.00%	0.00%	***0.10%***
	Overall	80.89%	79.67%	79.27%	79.27%	79.27%	***79.67%***
Recursive Partitioning	Type I error	50.98%	60.78%	90.20%	72.55%	68.63%	***68.63%***
	Type II error	5.64%	11.28%	7.69%	9.74%	5.13%	***7.90%***
	Overall	84.96%	78.46%	75.20%	77.24%	81.71%	***79.51%***

Table 8.6 Classification Accuracy Rates of the PCA Models (G12)

Methodology	Performance	M1	M2	M3	M4	M5	Average
Naïve Bayes	Type I error	35.29%	43.14%	54.90%	23.53%	27.45%	***36.86%***
	Type II error	7.18%	7.69%	7.69%	8.21%	7.18%	***7.59%***
	Overall	86.99%	84.96%	82.52%	88.62%	88.62%	***86.34%***
Logistic Regression	Type I error	43.14%	60.78%	60.78%	41.18%	54.90%	***52.16%***
	Type II error	3.59%	4.62%	3.08%	5.13%	6.15%	***4.51%***
	Overall	88.21%	83.74%	84.96%	87.40%	83.74%	***85.61%***

Table 8.6 Classification Accuracy Rates of the PCA Models (G12) (Continued.)

Methodology	Performance	M1	M2	M3	M4	M5	Average
Neural Network	Type I error	37.25%	47.06%	37.25%	37.25%	31.37%	***38.04%***
	Type II error	5.13%	5.13%	5.64%	5.64%	9.23%	***6.15%***
	Overall	88.21%	86.18%	87.80%	87.80%	86.18%	***87.23%***
SMO	Type I error	86.27%	84.31%	100.00%	62.75%	80.39%	***82.74%***
	Type II error	2.05%	1.03%	0.51%	2.56%	1.54%	***1.54%***
	Overall	80.49%	81.71%	78.86%	84.96%	82.11%	***81.63%***
Recursive Partitioning	Type I error	43.14%	43.14%	49.02%	31.37%	25.49%	***38.43%***
	Type II error	6.15%	4.62%	5.13%	4.10%	5.13%	***5.03%***
	Overall	86.18%	87.40%	85.77%	90.24%	90.65%	***88.05%***

8.3.1.1 Exploring Time Scale

As mentioned in Chapter Six, a five-year time scale analysis can be carried out by comparing the performance of models from five different time periods (M1, M2, M3, M4 and M5). M1 is designed for evaluating a model's performance one year before financial distress; M2 is designed for assessing a model's utility two years before financial distress and so on.

From Tables 8.5 and 8.6, almost all the credit-scoring techniques show best overall classification ability one year before financial distress. The exceptions are: Naïve Bayes model in G12 with the best overall classification performance four and five years before financial distress (M4 and M5), SMO model in G12 with the best overall classification performance four years before financial distress (M4), and Recursive Partitioning model in G12 with the best performance five years before financial distress (M5).

Regardless of the variable groups, the overall classification accuracy rate among different credit scoring techniques is over 80.49% in the year before financial distress (M1). Furthermore, even if the time period is five years before financial distress, almost all of the overall classification accuracy rates remain above 79.27%. Indeed, the only exception is the Naïve Bayes model in G1, which shows 65.04% overall accuracy rate five years before financial distress.

In terms of the average overall accuracy rate of five time scales, the Logistic Regression model and the Recursive Partitioning model show the best performance with the average accuracy rate of 83.49% in G1 and 88.05% in G12, respectively. However, the differences of the average accuracy rate among five credit scoring techniques in each variable group are small.

8.3.1.2 Types of Error

Tables 8.5 and 8.6 show that, in spite of different time scales, variable groups or credit scoring techniques, Type I error is always higher than the Type II error. The only exception is the Naïve Bayes model in G1, where Type II error (36.92%) is higher than Type I error (27.45%) five years before financial distress. Regardless of the variable group, the SMO model shows the best ability to deal with Type II error not only based on different time scale but also based on the average Type II error. However, the SMO model also presents the worst ability to cope with the Type I error vis-à-vis the different time periods or the average performance. A high Type I error also indicates that most sample companies are classified as healthy companies and it will damage the benefits from some interested parties. For example, Type I error may cause an investor to lose the entire investment, while Type II error may only cause an investor to lose the potential dividends or capital gains.

The Naïve Bayes model displays the best ability to manage Type I error apart from the variable group based on the average performance. Nevertheless, it is difficult to conclude which credit scoring model has the best performance to handle Type I error in terms of different time scales. For example, although Naïve Bayes model in G12 shows the best performance in the time periods M1 and M4, the same cannot be concluded in the time periods M2, M3 and M5.

8.3.1.3 Detecting External Influences

As mentioned in Chapter Six, external influences can be detected by comparing the performance of G1 and G12 models. If G12 performs better than G1, external factors have impacts on the model classification ability. The overall accuracy rate based on five credit scoring techniques among different time scales can be presented in Figures 8.2, 8.3, 8.4, 8.5 and 8.6 respectively.

Figure 8.2 Detecting External Influences: Naïve Bayes based on Accuracy Rate (PCA)

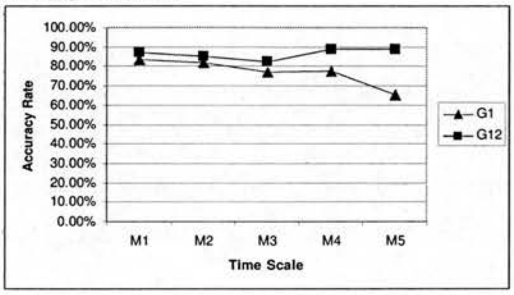


Figure 8.3 Detecting External Influences: Logistic Regression based on Accuracy Rate (PCA)

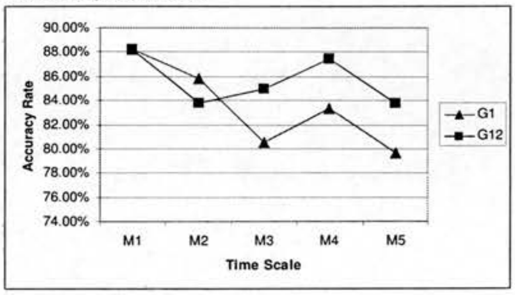


Figure 8.4 Detecting External Influences: Neural Network based on Accuracy Rate (PCA)

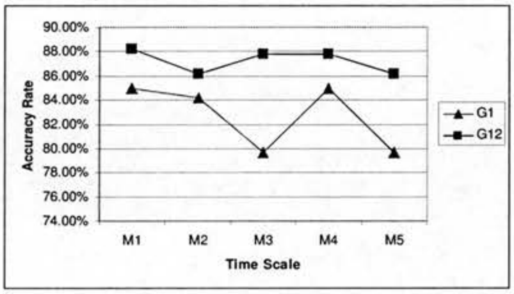


Figure 8.5 Detecting External Influences: SMO based on Accuracy Rate (PCA)

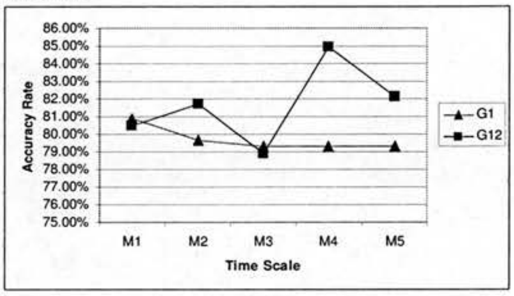
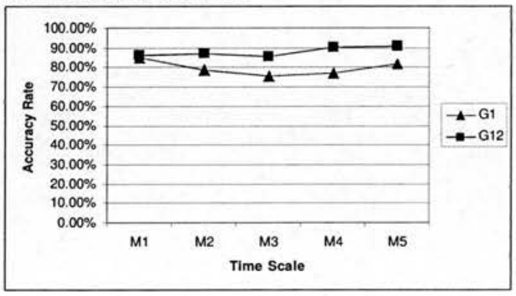


Figure 8.6 Detecting External Influences: Recursive Partitioning based on Accuracy Rate (PCA)



It is obvious from the figures that almost all G12 models display better accuracy rate through different time scales. The only special cases are the Logistic Regression model in M2 and SMO model in M3. The same can be concluded when comparing the average accuracy rate. The figures indicate that all G12 models perform better than G1 models in terms of the average five time scales performance. Moreover, the differences of the average accuracy rate based on the Naïve Bayes and Recursive Partitioning models are 9.43% and 8.54% respectively, whilst for the other three credit scoring models, the differences are below 5%. It can be concluded that the external influences have greater impacts on the Naïve Bayes and Recursive Partitioning models than other credit scoring PCA models.

8.3.2 AUROC Analysis

The results of the AUROC values and ROC curves are based on two variable groups in five different time scales. These are presented in Tables 8.7 and 8.8 as well as Figure 8.7 and 8.8 respectively:

Table 8.7 AUROC Values of the PCA Models (G1)

Methodology	M1	M2	M3	M4	M5	Average
Naïve Bayes	0.8456	0.7756	0.6750	0.7729	0.7345	***0.7607***
Logistic Regression	0.9063	0.8256	0.8046	0.8290	0.7896	***0.8310***
Neural Network	0.8815	0.8003	0.8271	0.8347	0.7954	***0.8278***

Figure 8.7 ROC Curves of the PCA Models (G1)

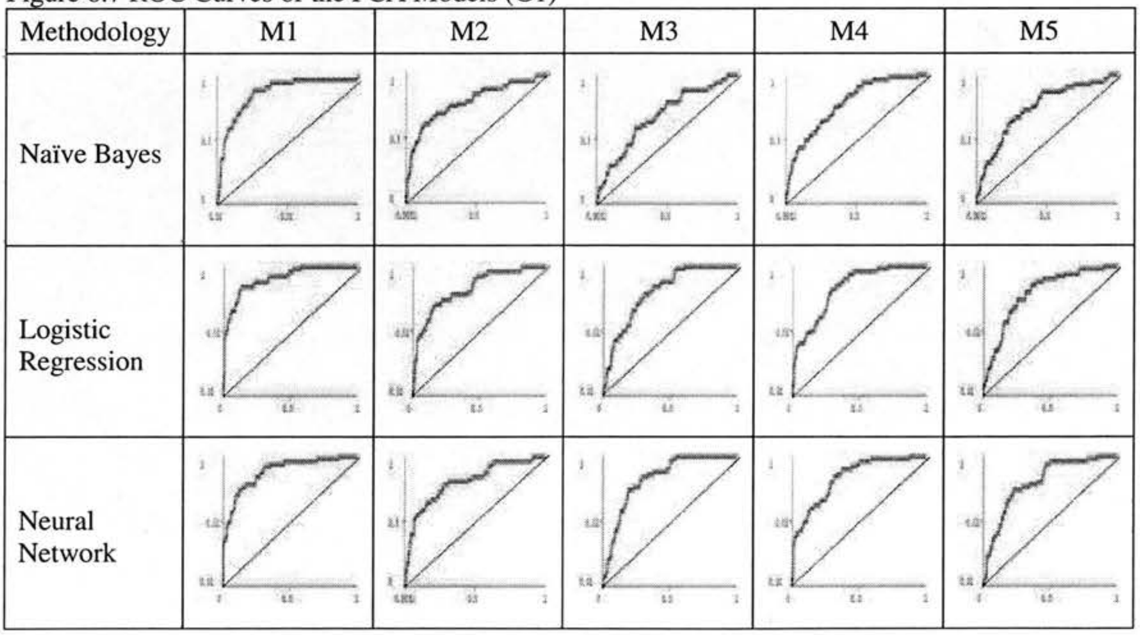
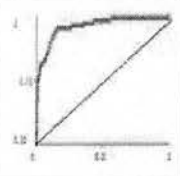
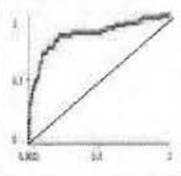
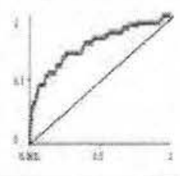
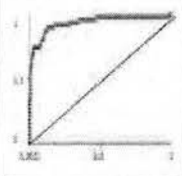
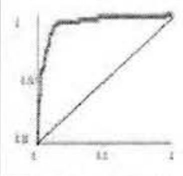
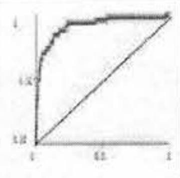
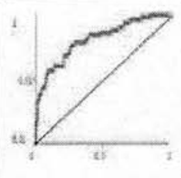
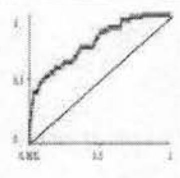
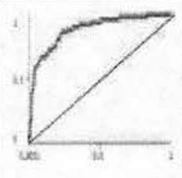
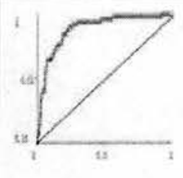
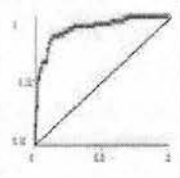
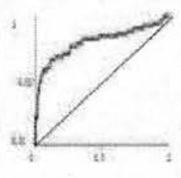
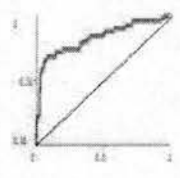
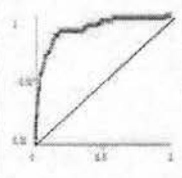
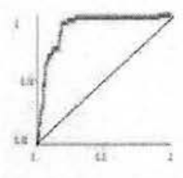


Table 8.8 AUROC Values of the PCA Models (G12)

Methodology	M1	M2	M3	M4	M5	Average
Naïve Bayes	0.9107	0.8470	0.7636	0.9320	0.9244	***0.8755***
Logistic Regression	0.9114	0.8055	0.7771	0.8800	0.8796	***0.8507***
Neural Network	0.9039	0.7977	0.8264	0.8982	0.8982	***0.8649***

Figure 8.8 ROC Curves of the PCA Models (G12)

Methodology	M1	M2	M3	M4	M5
Naïve Bayes					
Logistic Regression					
Neural Network					

8.3.2.1 Exploring Time Scale

Tables 8.7 and 8.8 show that almost all the credit scoring models perform best in the M1 time period. As with the accuracy rate analysis, the exception is the Naïve Bayes model in the time scale of M4 in G12. In addition, regardless of the variable group, the AUROC values in the year before financial distress (M1) are above 0.8456. This implies that all credit scoring models present sound discriminating performance. Furthermore, regardless of the variable group, in the time period of five years prior to default, the AUROC value is still higher than 0.7345. As a result, based on the definition in Hosmer and Lemeshow (2000), the discriminating power is still acceptable in relation to long-term prediction performance.

However, it is difficult to conclude which modelling technique has the *'absolute'* best performance, since the model's AUROC values vary in terms of different time scales and variable groups. For example, Logistic Regression model shows the best performance in M1, but the same cannot be concluded in different time periods. The results can also be detected from Figure 8.8. For example, Naïve Bayes model shows the largest AUROC area in M2, but Neural Network model presents the largest AUROC area in M3. Based on the average AUROC value, the Naïve Bayes displays the best performance in G12, but shows thc worst performance in G1.

8.3.2.2 Detecting External Influences

Similar to Section 8.3.1.3, the following line charts (Figure 8.9, 8.10 and 8.11) can be employed to detect the existence of external influences:

Figure 8.9 Detecting External Influences: Naïve Bayes based on the AUROC (PCA)

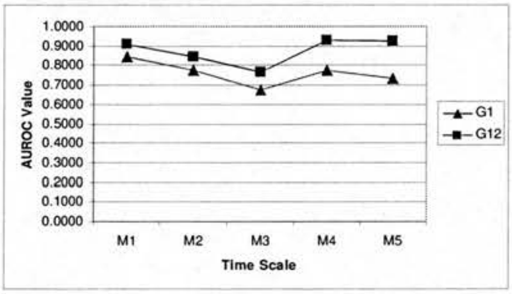


Figure 8.10 Detecting External Influences: Logistic Regression based on the AUROC (PCA)

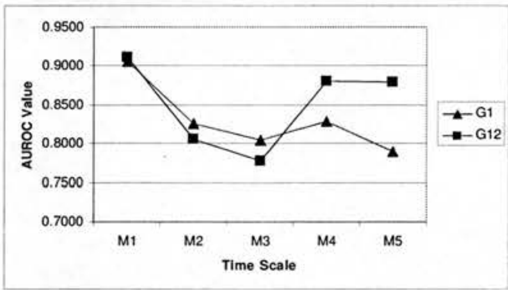
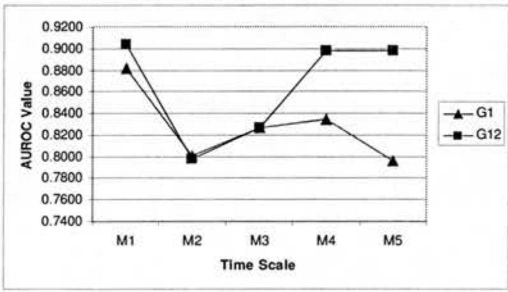


Figure 8.11 Detecting External Influences: Neural Network based on the AUROC (PCA)



They show that external factors are more important in Naïve Bayes model than in the other two credit scoring models, since G12 performs better than G1 in all five time scales. The same can be concluded by looking at the difference of average

AUROC between G1 and G12. The difference for Naïve Bayes model is 0.1148, but for Logistic Regression and Neural Network models are 0.0197 and 0.0371 respectively.

8.3.3 Concluding Remarks for PCA Model Analysis

Regarding the exploring of time scale and variable groups, almost all credit-scoring models show best performance in the year before financial distress (with the accuracy rate of above 80.49% and AUROC value of above 0.8456). In the time period of five years prior to financial distress, the accuracy rate is above 79.27% and the AUROC value is above 0.7345. (The only exception is the Naïve Bayes model in G1, which shows an accuracy rate of 65.04% in M5) However, it is difficult to conclude which credit scoring technique has the *'absolute'* best performance, since the model's utility varies in terms of different time scales and variable groups.

On the subject of types of error, Type I error is higher than Type II error in most cases. The only exception is the Naïve Bayes model in G1, five years before financial distress. A high Type I error also indicates that most sample companies are classified as healthy companies. SMO model shows the best ability to deal with Type II error, whilst it also presents the worst ability to cope with Type I error. Naïve Bayes model displays the best ability to manage Type I error apart from the variable group based on the average performance. However, the same cannot be concluded in terms of different time periods, since Naïve Bayes model does not always show the best performance in each time period.

Turning to external influences, it can be concluded that these exist in terms of both the accuracy rate and the AUROC value analyses, as the average performance in G12 is better than the average performance in G1 among almost all five credit scoring models. However, the importance of external influences is different among the credit scoring models. The results indicate that the external influences have greater impact on Naïve Bayes and Recursive Partitioning models than other PCA credit scoring models.

8.4 Prediction Utility Assessment for Stepwise Regression Models

8.4.1 Classification Accuracy Rate Analysis

The results of the classification accuracy rates based on two variable groups (G1 and G12) among five different time scales are displayed in Tables 8.9 and 8.10:

Table 8.9 Classification Accuracy Rates of the Stepwise Regression Models (G1)

Methodology	Performance	M1	M2	M3	M4	M5	Average
Naïve Bayes	Type I error	35.29%	39.22%	50.98%	54.90%	52.94%	***46.67%***
	Type II error	4.62%	9.23%	10.77%	9.23%	9.74%	***8.72%***
	Overall	89.02%	84.55%	80.89%	81.30%	81.30%	***83.41%***
Logistic Regression	Type I error	29.41%	45.10%	64.71%	68.63%	70.59%	***55.69%***
	Type II error	5.13%	4.62%	6.15%	4.62%	5.64%	***5.23%***
	Overall	89.84%	86.99%	81.71%	82.11%	80.89%	***84.31%***
Neural Network	Type I error	25.49%	29.41%	45.10%	60.78%	52.94%	***42.74%***
	Type II error	2.05%	3.59%	4.10%	5.64%	2.56%	***3.59%***
	Overall	93.09%	91.06%	87.40%	82.93%	86.99%	***88.29%***
SMO	Type I error	45.10%	45.10%	56.86%	82.35%	100.00%	***65.88%***
	Type II error	1.03%	2.05%	2.56%	1.03%	1.03%	***1.54%***
	Overall	89.84%	89.02%	86.18%	82.11%	78.46%	***85.12%***
Recursive Partitioning	Type I error	21.57%	35.29%	37.25%	62.75%	54.90%	***42.35%***
	Type II error	4.10%	5.64%	4.10%	1.54%	3.59%	***3.79%***
	Overall	92.28%	88.21%	89.02%	85.77%	85.77%	***88.21%***

Table 8.10 Classification Accuracy Rates of the Stepwise Regression Models (G12)

Methodology	Performance	M1	M2	M3	M4	M5	Average
Naïve Bayes	Type I error	15.69%	25.49%	27.45%	29.41%	23.53%	***24.31%***
	Type II error	7.18%	7.69%	8.21%	8.72%	8.21%	***8.00%***
	Overall	91.06%	88.62%	87.80%	86.99%	88.62%	***88.62%***
Logistic Regression	Type I error	27.45%	37.25%	43.14%	45.10%	43.14%	***39.22%***
	Type II error	3.08%	3.59%	3.59%	5.13%	4.62%	***4.00%***
	Overall	91.87%	89.43%	88.21%	86.59%	87.40%	***88.70%***
Neural Network	Type I error	23.53%	35.29%	33.33%	35.29%	27.45%	***30.98%***
	Type II error	6.15%	4.10%	6.67%	6.15%	7.69%	***6.15%***
	Overall	90.24%	89.43%	87.80%	87.80%	88.21%	***88.70%***

Table 8.10 Classification Accuracy Rates of the Stepwise Regression Models (G12) (Continued.)

SMO	Type I error	37.25%	41.18%	54.90%	52.94%	45.10%	***46.27%***
	Type II error	2.56%	2.56%	3.59%	4.10%	4.10%	***3.38%***
	Overall	90.24%	89.43%	85.77%	85.77%	87.40%	***87.72%***
Recursive Partitioning	Type I error	19.61%	33.33%	29.41%	37.25%	35.29%	***30.98%***
	Type II error	4.10%	4.62%	7.18%	5.13%	4.62%	***5.13%***
	Overall	92.68%	89.43%	88.21%	88.21%	89.02%	***89.51%***

8.4.1.1 Exploring Time Scale

Tables 8.9 and 8.10 show that all credit-scoring models are overall most accurate one year prior to default (M1). Regardless the variable groups, the accuracy rates in M1 are above 89.02%. Even five years before financial distress, the accuracy rate is still above 78.46%. This suggests that the overall performance of these five modelling methodologies is sound, even if the time period chosen is as long as five years before financial distress. Furthermore, these results prove that the five key variables selected are effective for financial distress predictions.

Regarding average performance, the Neural Network model shows the best performance in G1 with an average accuracy rate of 88.29%. Recursive Partitioning model presents the best performance in G12 with the average accuracy rate of 89.51%. However, this does not mean that the Neural Network model and the Recursive Partitioning model have the highest accuracy rate in terms of different time periods in G1 and G12 respectively. For example, in G12, Recursive Partitioning model only shows best performance in M1, M4 and M5. For the other two time periods, it shares the same performance with other credit scoring models.

8.4.1.2 Types of Error

From Tables 8.9 and 8.10, it is clear that Type II error is smaller than Type I error for all variable groups, time scales and modelling techniques. As with PCA model analysis, SMO model in both variable groups shows the best ability to deal with the Type II error, but also displays the worst ability to cope with the Type I error.

With regards to the ability to control Type I error, Naïve Bayes model shows the best performance based on the average Type I error in G12, whilst Recursive Partitioning model presents the best performance in G1. This conclusion is slightly different from the PCA model analysis, where Naïve Bayes PCA model shows best performance based on the average Type I error in G1.

8.4.1.3 Detecting External Influences

As in previous analyses, line charts are employed here to discuss the existence of external influences. (see Figures 8.12, 8.13, 8.14, 8.15 and 8.16.)

Figure 8.12 Detecting External Influences: Naïve Bayes based on Accuracy Rate (Stepwise)

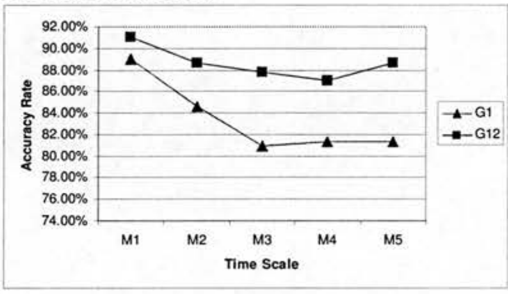


Figure 8.13 Detecting External Influences: Logistic Regression based on Accuracy Rate (Stepwise)

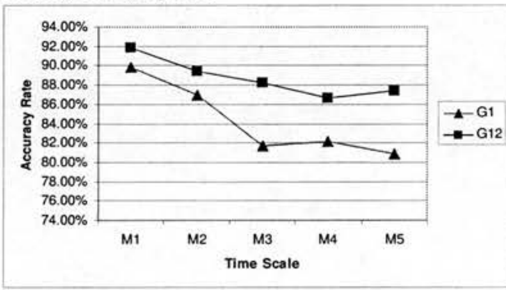


Figure 8.14 Detecting External Influences: Neural Network based on Accuracy Rate (Stepwise)

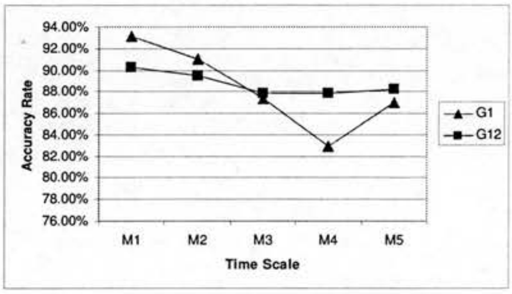


Figure 8.15 Detecting External Influences: SMO based on Accuracy Rate (Stepwise)

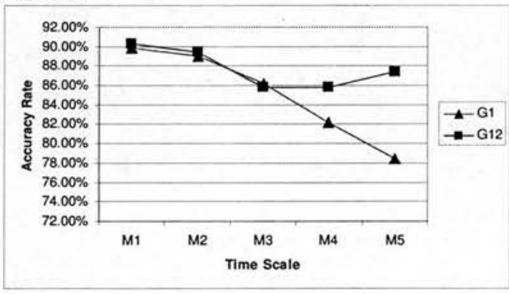
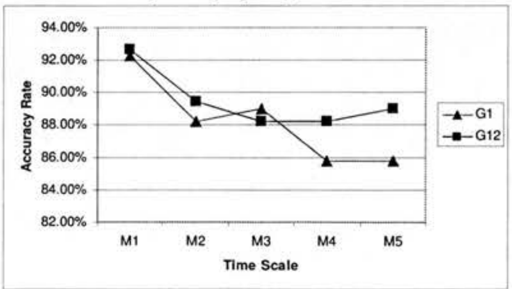


Figure 8.16 Detecting External Influences: Recursive Partitioning based on Accuracy Rate (Stepwise)



Figures 8.12 to 8.16 demonstrate better accuracy rate for almost all G12 models than G1 models in different time periods. The only special cases are the Neural Network model in M1 and M2, SMO model in M3, and Recursive Partitioning model in M3. When comparing average accuracy rates, all G12 models also perform better than G1 models each time. However, unlike the PCA model analysis, the differences of the average accuracy rate among five credit-scoring are small (all below 5.21%). It can be concluded that external environment influences exist in all modelling methodologies, but these influences are weak.

8.4.2 AUROC Analysis

AUROC values and ROC curves of the two variable groups in five time scales are presented in Tables 8.11 and 8.12 as well as Figures 8.17 and 8.18:

Table 8.11 AUROC Values of the Stepwise Regression Models (G1)

Methodology	M1	M2	M3	M4	M5	Average
Naïve Bayes	0.9161	0.8792	0.8155	0.7798	0.8140	***0.8409***
Logistic Regression	0.9341	0.8860	0.8156	0.7816	0.7955	***0.8426***
Neural Network	0.9158	0.9024	0.8498	0.7982	0.8755	***0.8683***

Figure 8.17 ROC Curves of the Stepwise Regression Models (G1)

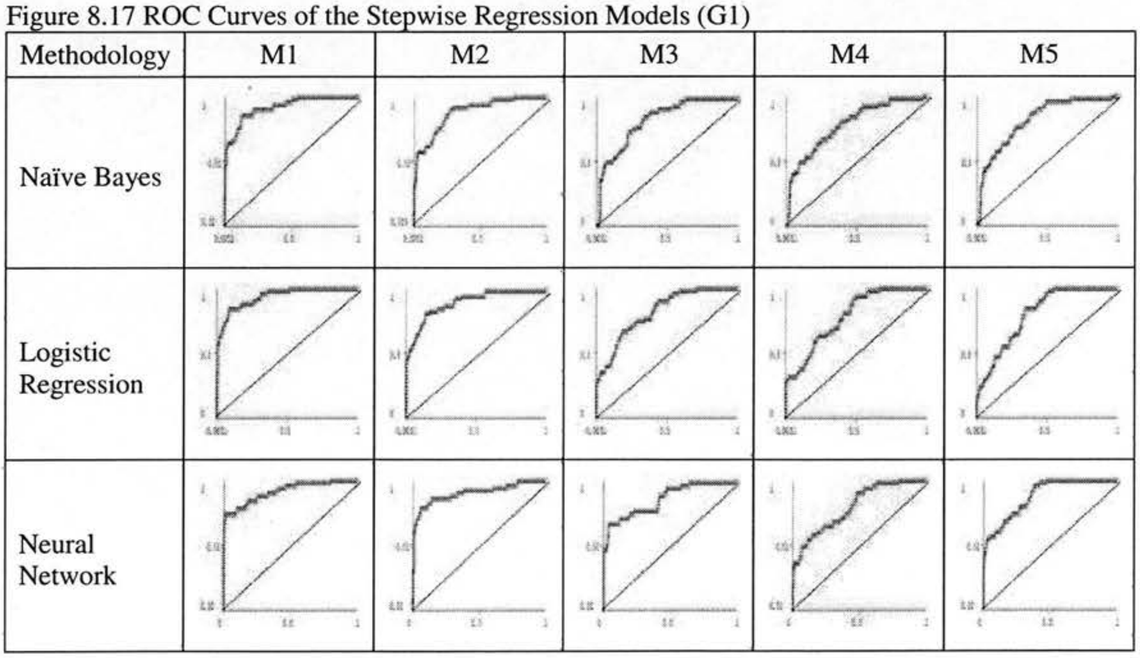
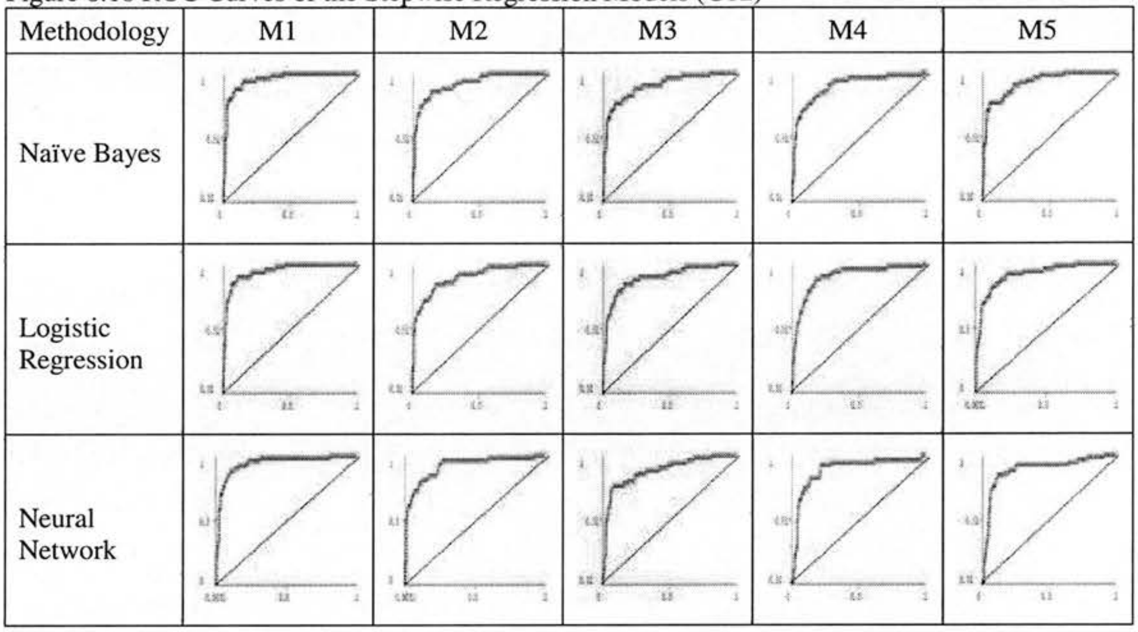


Table 8.12 AUROC Values of the Stepwise Regression Models (G12)

Methodology	M1	M2	M3	M4	M5	Average
Naïve Bayes	0.9509	0.9174	0.8967	0.8950	0.9158	***0.9152***
Logistic Regression	0.9448	0.8970	0.8894	0.8964	0.9079	***0.9071***
Neural Network	0.9350	0.9140	0.8762	0.8808	0.8794	***0.8971***

Figure 8.18 ROC Curves of the Stepwise Regression Models (G12)



8.4.2.1 Exploring Time Scale

Tables 8.11 and 8.12 as well as Figures 8.17 and 8.18 brings forward the same conclusion as in previous analyses. All credit scoring models show the best accuracy rate in the time scale of one year before financial distress (all AUROC values are above 0.9158). In addition, even if the time period is five years before financial distress, the AUROC value is above 0.7955. The results again suggest that the overall performance of the five modelling methodologies is sound, even in a long time period. Moreover, it also indicates that the five key variables are effective for predicting financial distress.

With regards to average performance, Neural Network shows the best performance in G1 and the worst performance in G12. Moreover, the performance of each credit scoring model fluctuates in different time periods. For example, the Logistic Regression model in G12 displays the best performance in the year prior to financial distress (M1), but the same cannot be concluded in different time periods. As with PCA analysis, it is difficult to conclude which credit scoring technique has the *'absolute'* best performance.

8.4.2.2 Detecting External Influences

Figures 8.19, 8.20 and 8.21 are line charts illustrating external environmental impacts:

Figure 8.19 Detecting External Influences: Naïve Bayes based on AUROC Analysis (Stepwise)

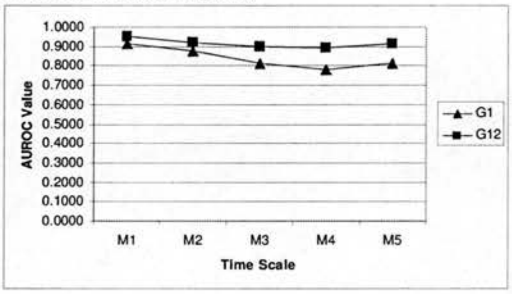


Figure 8.20 Detecting External Influences: Logistic Regression based on AUROC Analysis (Stepwise)

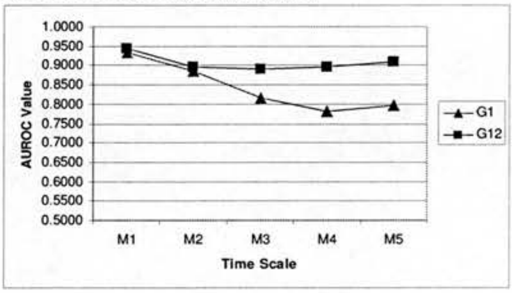
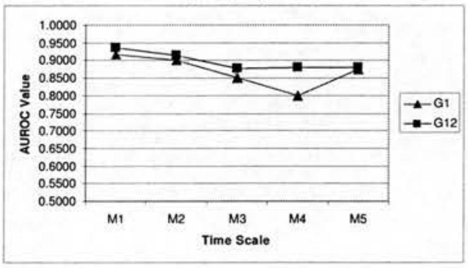


Figure 8.21 Detecting External Influences: Neural Network based on AUROC Analysis (Stepwise)



In the line charts above, all G12 models show better performance than G1 models in different time periods. In terms of average AUROC values, all G12 models display higher average AUROC values than G1 models. However, the difference of the average AUROC value between G1 and G12 is small (below 0.0743) for all three credit scoring techniques. As a result, the same conclusion as that of accuracy rate analysis can be reached: external environment influences exist in all modelling methodologies, but these influences are weak.

8.4.3 Concluding Remarks for Stepwise Regression Model Analysis

Regarding the issue of exploring time scale, all credit scoring models in G1 and G12 show the best performance in the year prior to default, with the accuracy rate of above 89.02% and AUROC value of above 0.9158. Moreover, in the time period of five years prior to financial distress, the accuracy rate is above 78.46% and the AUROC value is above 0.7955. The results suggest that the overall performance of these five modelling methodologies is sound, even if the time period chosen is as long as five years before financial distress. Furthermore, they also prove that the key variables selected are effective for predicting financial distress. However, it is difficult to conclude which credit scoring technique has the *'absolute'* best performance, since model utility varies depending on different time scales and variable groups.

On the topic of types of error, Type I error is higher than the Type II error regardless of time scales, variable groups or the credit scoring techniques. SMO results are similar to PCA's: both have the best ability to deal with the Type II error, but the worst ability when coping with the Type I error. Regarding the ability to deal with the Type I error, Naïve Bayes model shows the best performance based on the average Type I error in G12, whilst Recursive Partitioning model presents the best performance in G1.

With regards to the detection of the external influences, it can be concluded that the external influences exist in both the accuracy rate and the AUROC value analyses. However, as the difference in average accuracy rate and AUROC value between G1 and G12 is small, the external influences are weak.

Thus far, this research has carried out an evaluation of the prediction power in terms of both PCA and Forward Stepwise models. The following section presents a comparative analysis between these two variable selection approaches.

8.5 Comparative Analysis Between PCA and Stepwise Regression Models

8.5.1 Accuracy Rate Comparative Analysis

The following line charts (from Figure 8.22 to Figure 8.31) can be employed to facilitate the accuracy rate comparative analysis in different time scales.

Figure 8.22 Naïve Bayes Model Comparative Analysis based on Accuracy Rate (G1)

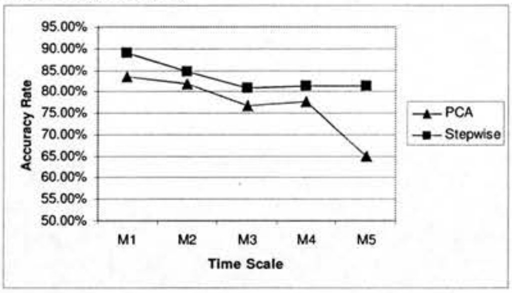


Figure 8.23 Logistic Regression Model Comparative Analysis based on Accuracy Rate (G1)

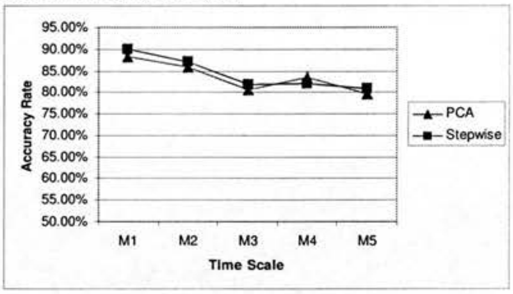


Figure 8.24 Neural Network Model Comparative Analysis based on Accuracy Rate (G1)

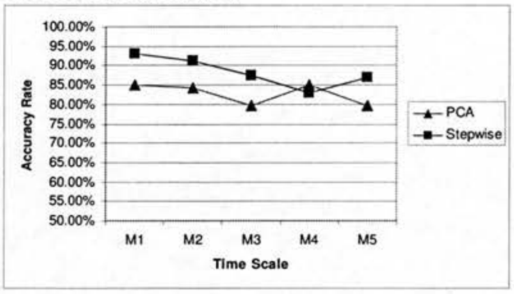


Figure 8.25 SMO Model Comparative Analysis based on Accuracy Rate (G1)

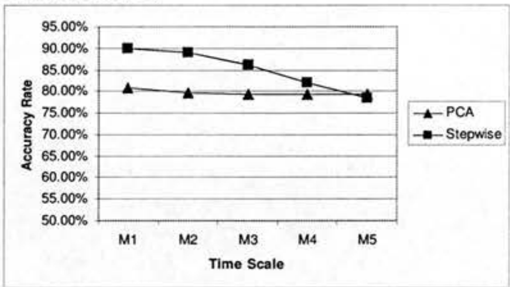


Figure 8.26 Recursive Partitioning Model Comparative Analysis based on Accuracy Rate (G1)

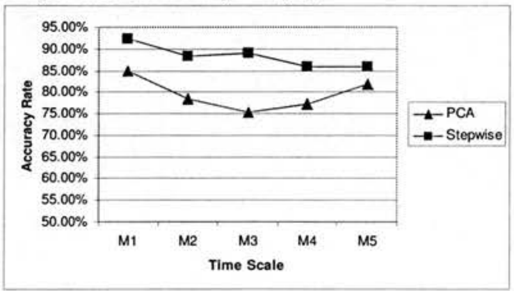


Figure 8.27 Naïve Bayes Model Comparative Analysis based on Accuracy Rate (G12)

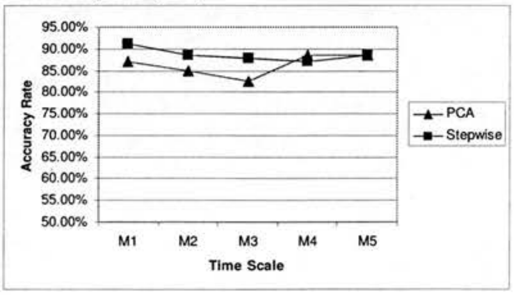


Figure 8.28 Logistic Regression Model Comparative Analysis based on Accuracy Rate (G12)

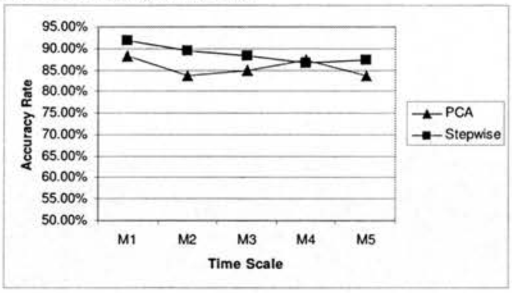


Figure 8.29 Neural Network Model Comparative Analysis based on Accuracy Rate (G12)

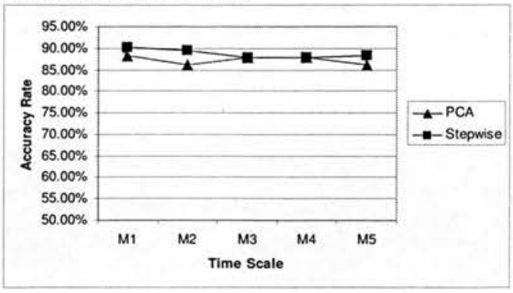


Figure 8.30 SMO Model Comparative Analysis based on Accuracy Rate (G12)

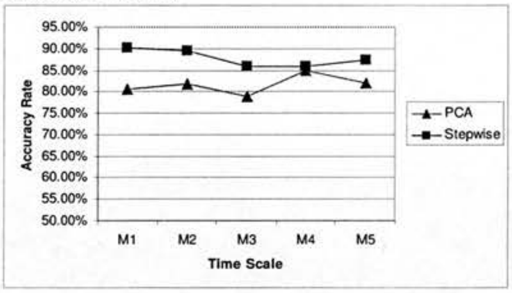
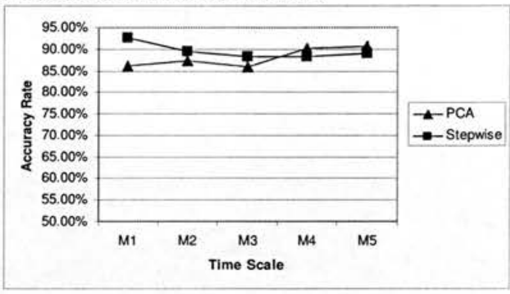


Figure 8.31 Recursive Partitioning Model Comparative Analysis based on Accuracy Rate (G12)



Figures 8.22 to 8.31 show that almost all the Forward Stepwise models have higher accuracy rate than PCA models for both G1 and G12. The special cases are: Logistic Regression model in G1 in M4, Neural Network model in G1 in M4, SMO model in G1 in M5, Naïve Bayes model in G12 in M4, Logistic Regression model in G12 in M4, and Recursive Partitioning model in G12 in M4 and M5. A comparative analysis in terms of the average accuracy rate confirms the above: all Forward Stepwise models show higher average accuracy rate than the PCA model.

In addition, the accuracy rate of the PCA models and the Forward Stepwise models tend to converge in the time period of four or five years prior to financial distress. In other words, the longer the time period before financial distress, the smaller the difference in performance between the PCA models and Forward Stepwise models is. Therefore, it can be concluded that the superior performance of Forward Stepwise models in comparison with the PCA models is only obvious in the time periods one, two or three years prior to financial distress.

8.5.2 AUROC Value Comparative Analysis

As in accuracy rate analysis, the following line charts (from Figure 8.32 to Figure 8.37) are used here to carry out the AUROC comparative analysis between the PCA models and the Forward Stepwise models.

Figure 8.32 Naïve Bayes Model Comparative Analysis based on AUROC (G1)

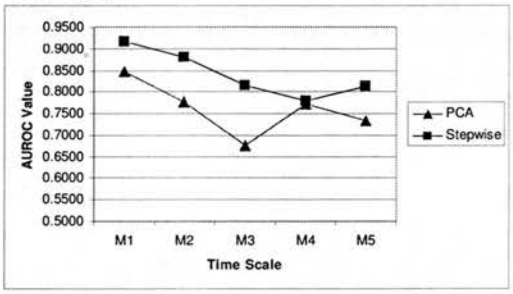


Figure 8.33 Logistic Regression Model Comparative Analysis based on AUROC (G1)

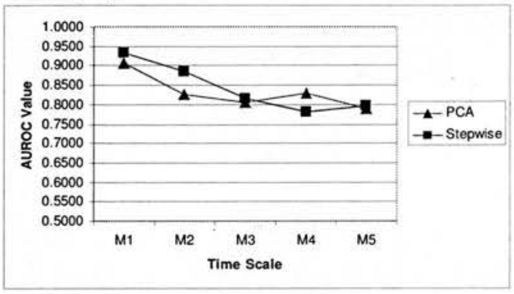


Figure 8.34 Neural Network Model Comparative Analysis based on AUROC (G1)

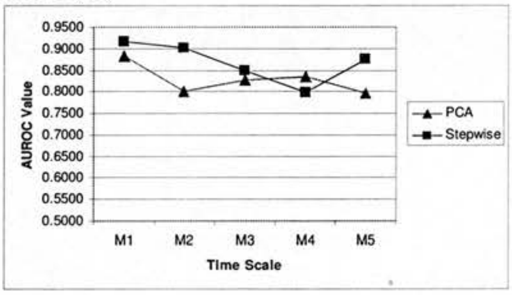


Figure 8.35 Naïve Bayes Model Comparative Analysis based on AUROC (G12)

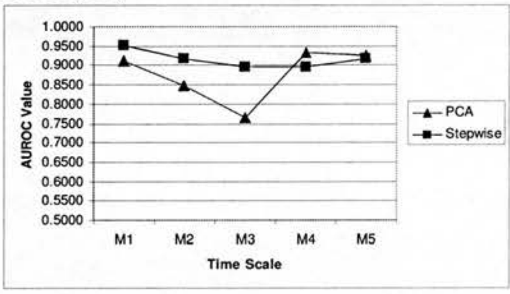


Figure 8.36 Logistic Regression Model Comparative Analysis based on AUROC (G12)

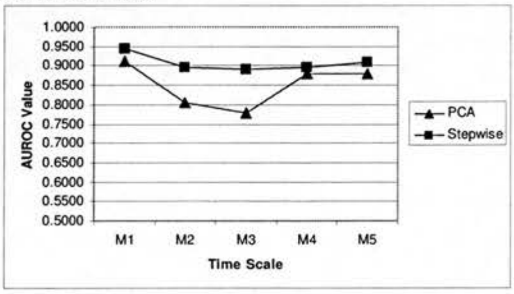
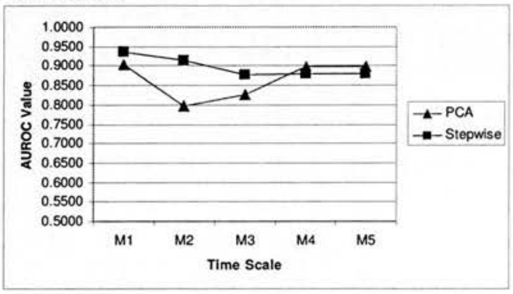


Figure 8.37 Neural Network Model Comparative Analysis based on AUROC (G12)



AUROC analysis display results similar to accuracy rate analysis. Regardless the variable groups, almost all Forward Stepwise models show higher AUROC values than PCA model. The same applies to the average AUROC values. Finally, Forward Stepwise models continue to display better performance than PCA models, especially in the time periods one, two or three years before financial distress.

Drawing on above, the Forward Stepwise models have better classification performance than PCA models in terms of both average accuracy rate and average AUROC value. A possible reason for this finding is that the theoretical foundation of variable selection criterion is different between the Forward Stepwise approach and the PCA approach. Indeed, Forward Stepwise approach uses discriminating performance as the main criterion. In other words, final variables selected in the Forward Stepwise model must have basic classification power. However, in the PCA model, the principal components are selected in terms of their ability to explain total variance. As a result, these principal components do not have identical classification power as the variables in the Forward Stepwise model.

Another possible reason is that only the first five principal components are selected for the model construction purposes. With a total explained variance ranging from 55% to 65%, the PCA models actually ignore approximately 40% of total variance. This leads to a worse classification utility. If more explained variance were considered for final principal components selection, results may differ.

8.6 Concluding Remarks

This chapter evaluated the prediction performance of the credit scoring models based on two variable selection methods: Principal Component Analysis (PCA) and Forward Stepwise Approach. The classification accuracy rate and AUROC value were employed in order to discuss three key issues: exploring time scale, types of error and the detecting external influences.

Regarding performance by different pre-distress time scales, almost all credit scoring models displayed the best performance in the year prior to financial distress. In the year before default, PCA models show an accuracy rate of above 80.49% and AUROC value above 0.8456. For Forward Stepwise models, the accuracy rate is above 89.02% and AUROC value is above 0.9158. Five years before financial distress, the accuracy rate and AUROC value remain high: above 79.27% and above 0.7345 respectively for PCA models, or above 78.46% and above 0.7955 respectively for Forward Stepwise models. Such results suggest that the overall performance of these five modelling methodologies is sound, even if the time period chosen is as long as five years before financial distress. Furthermore, the key variables selected are effective for predicting financial distress. However, it is difficult to conclude which credit scoring technique has the *'absolute'* best performance, since the model's utility varies in terms of different time scales and variable groups.

For types of error, the Type I error is higher than the Type II error in all cases except for the Naïve Bayes PCA model in G1 in the time period of five years before financial distress. A high Type I error also indicates that most sample companies are classified as healthy companies and it will damage the benefits from some interested parties. For example, Type I error may cause an investor to lose the entire investment, while Type II error may only cause an investor to lose the potential dividends or capital gains.

The study also shows that SMO model has the best ability to deal with the Type II error, whilst it also presents the worst ability to cope with the Type I error. With regards to the Type I error control ability, the Naïve Bayes PCA model displays the best ability to manage Type I error based on the average performance. However, the same cannot be concluded for the Forward Stepwise models, as Naïve Bayes model only shows the best performance based on the average Type I error in G12. In G1, the Recursive Partitioning model presents the best performance to dealing with Type I error.

On the topic of the detecting external influences, the external influences exist in terms of both the accuracy rate and the AUROC value analyses, as the average performance in G12 is better than the average performance in G1 among all five credit scoring models. For PCA models, the results indicated that external influences have greater impacts on Naïve Bayes and Recursive Partitioning models than on other PCA credit scoring models. However, for Forward Stepwise models, as the difference of the average accuracy rate and AUROC value between G1 and G12 is small among all credit coring techniques, it can be concluded that external influences are weak when stepwise regression approach is employed.

Finally, a comparative analysis between PCA models and Forward Stepwise models was conducted. Results showed that almost all Forward Stepwise models possess higher accuracy rate and AUROC values than PCA models. This is more obvious in one, two and three years before default. A possible reason for this result is that the variable selection criterion for stepwise approach is based on the variable's classification power, whereas that for PCA approach is based on the ability to explain total variance. Also, only the first five principal components were selected to construct prediction models. The explained variance is only 60% approximately. Hence, the findings may change if more explained variance were considered for final principal components selection.

Chapter Eight has assessed the model prediction performance in terms of both PCA and Forward Stepwise approaches with results indicating that all the credit scoring models perform best one year before financial distress. Having said this, all credit scoring models still remain sound five years prior to financial distress. Given the size of the sample for study it was not possible, and probably it would not have been informative, to employ a hold out sample. The above findings may result in potentially overly optimistic conclusions.

To overcome this problem, the researcher decided to compare the results from the study with a standard rating system such as Moody's rating. The main reason for comparing the research models with Moody's rating is the lack of hold out sample to give a cross-validation of the credit scoring models especially in a practical context. The results on the research models are not exactly the same as Moody's. Hence, some benefits may be gained from these new models. Further investigation is worthwhile, especially for SMO, which has not been considered in previous works. As most G12 models performed better than G1 models, and most credit scoring models showed the best performance in M1, Logistic Regression, Neural Network and SMO models in the time period one year before financial distress in the variable group G12 were selected for the ranking comparison analysis. The issues of the model's practical applicability will be discussed in the next chapter.

Chapter NINE

Model Practical Applicability Evaluation: Comparison with Moody's Rating

9.1 Introduction

In Chapter Eight, the researcher examined the prediction utility of the credit scoring models for both PCA and Forward Stepwise approaches. Results indicated that all credit-scoring models fitted to the data performed well from one to five years before financial distress. Since the size of dataset did not allow a hold out sample, it was felt that a comparison should be made with an alternative external rating, and Moody's rating was chosen. The comparison to Moody's rating will be the purpose of this chapter.

Six credit scoring models from the research, which include three PCA models and three Forward Stepwise models, were selected for the comparative analysis with Moody's rating. Section 9.3 then describes the Moody's rating data collection. The approaches for comparative analysis will be introduced in the Section 9.4. The comparative analyses with Moody's rating in terms of both PCA models and Forward Stepwise models will be presented in the Sections 9.5 and 9.6 respectively. Section 9.7 compares the results from PCA models and Forward Stepwise models. The final section summarises the main findings in this chapter.

9.2 Credit Scoring Models for Practical Applicability Evaluation

In the previous chapter, two main conclusions were reached: first, for all variable selection approaches, almost all credit scoring models showed the best prediction performance one year before financial distress (M1); second, the external influences exist among all credit scoring models, since G12 models displayed better performance than G1 models. Drawing on these two findings, it can be argued that the models in the time period one year before financial distress (M1) in the variable group G12 have better prediction utility. Hence, Logistic Regression, Neural Network and SMO models in the time period one year before financial distress in the variable group G12 are selected for ranking comparisons. As this research employed two variable selection approaches, overall six credit scoring models were selected for comparing with Moody's.

Credit-scoring analyses produce scores that are related to the companies' probability of default or bankruptcy and thus achieve the goal of predicting corporate performance. Based on the credit scores, each sample company will be attributed a credit rating. The rating data will then be compared with Moody's ratings in order to examine the practical applicability of the six credit scoring models.

9.3 Moody's Rating

Traditionally, Moody's[1] employs nine ranking categories from C (the lowest ranking) to Aaa (the highest ranking) to evaluate a company's long-term credit situation. For example, the grade of Aaa implies that a company's interest payments are protected by a large stable margin and any extra influence is unlikely to waver its strong position. In contrast, grade C means that a company is in default and the potential recovery values are low.

In the US retail industry, there are only 8 rating grades given in Moody's system (Aa to C). Therefore, in this study, rating data is ranked according to credit score and also divided into 8 groups with the same sample size. Unfortunately, Moody's ratings were only available for a limited number of retail companies, since firms undergo the credit rating process due to special circumstances, such as issuing corporate bond. Therefore, the sample size for analysis varies year on year. The sample size in different time periods is presented in Table 9.1:

Table 9.1 Sample Size of Rating Data

Time Period	2002	2001	2000	1999	1998
Sample size	72	73	75	77	73

1 Information is available at: http://www.moodys.com/

9.4 Techniques for Comparative Analysis

Four techniques—Kolmogorov-Smirnov (K-S) test, Distance analysis, Weighted Kappa analysis and Graphical Bubble charts—were used for comparative analysis. Each will be discussed in turn.

9.4.1 Kolmogorov-Smirnov (K-S) Test

Kolmogorov-Smirnov (K-S) test, as introduced in the Chapter Five for e-questionnaire comparative analysis, is used to investigate the significance of difference between two independent sample distributions. It first determines the cumulative distribution functions of both rating samples and then calculates the maximum absolute difference between two cumulative distribution functions. The basic rule is that if the maximum absolute difference is significantly large, then the two distributions are considered different. Therefore, if the *p*-value is greater than 0.05, then the two samples are likely to belong to the same distribution function. In this research, K-S significance testing is used to determine whether or not there is similarity in ranking. Distance analysis, Weighted Kappa analysis and Graphical Bubble chart techniques then attempt to assess the degree of similarity.

9.4.2 Distance Analysis

The most straightforward approach for analyzing the degree of similarity between two ordinal data sets is distance analysis. The rule is: the smaller the distance between the rankings from Moody's and those from the present study, the better the practical applicability of the present study's proposed model. In an 8×8 crosstabulation table (see Table 9.2), the diagonal depicts perfect match between the rankings and the distance is, therefore, zero. The other cells show distances from the diagonal line. For example, one cell away from the diagonal will be given a distance of 1. Two cells away from the diagonal will be given a distance of 2 and so forth.

Table 9.2 Distance Matrix Table

		Credit Scoring Model Rating Result							
		Rank1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	Rank 7	Rank 8
Moody's Rating Result	Rank 1	0	1	2	3	4	5	6	7
	Rank 2	1	0	1	2	3	4	5	6
	Rank 3	2	1	0	1	2	3	4	5
	Rank 4	3	2	1	0	1	2	3	4
	Rank 5	4	3	2	1	0	1	2	3
	Rank 6	5	4	3	2	1	0	1	2
	Rank 7	6	5	4	3	2	1	0	1
	Rank 8	7	6	5	4	3	2	1	0

To calculate distances, each cell is presented as a proportion of the total sample size. (This allows for year-on-year comparison, as the sample size of each year is different.) The cell value is then multiplied by the value in the distance matrix (Table 9.2). Finally, the resulting values are summed up. If the sum of distances is high, then the degree of similarity between Moody's rating and research model's rating is low. Low similarity can imply less practical applicability of the present study's proposed model, and vice versa for high similarity.

9.4.3 Weighted Kappa

When companies are evaluated by different raters, it is important to measure the degree of agreement between these raters. How much do the ratings provided by the Logistic Regression model, the Neural Network model, and the SMO model concord with those from Moody's? To answer this question, weighted Kappa was used. Cohen's Kappa (1960) is a measure of agreement between different raters only suitable for nominal data. Weighted Kappa is an extension of Cohen's Kappa suitable for ordinal data (as in ranking data) and for measuring relative concordance.

In an 8 × 8 crosstabulation rating table, each cell can be presented as n_{ij}, where i is the Moody's rating and j is the rating of the research model and the total sample size is N. For example, n_{23} indicates that the Moody's rating is 2 and the rating of the research model is 3. Clearly, the cells in the diagonal line, (such as n_{11}, n_{22}...) reflect the perfect match. In order to calculate weighted Kappa, each cell is attributed a weight w_{ij}. The weight in the Weight Matrix table is calculated from Table 9.2. Let each cell in the Distance Matrix table is D_{ij}. w_{ij} is calculated by the function:

$$w_{ij} = 1 - \frac{D_{ij}}{7} \tag{9.1}$$

The Weight Matrix can be expressed in Table 9.3. In Table 9.3, the weight *'1'* means totally match whilst the weight *'0'* means the lowest similarity.

Table 9.3 Weight Matrix Table

		Credit Scoring Model Rating Result							
		Rank1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	Rank 7	Rank 8
Moody's Rating Result	Rank 1	1.0000	0.8571	0.7143	0.5714	0.4286	0.2857	0.1429	0.0000
	Rank 2	0.8571	1.0000	0.8571	0.7143	0.5714	0.4286	0.2857	0.1429
	Rank 3	0.7143	0.8571	1.0000	0.8571	0.7143	0.5714	0.4286	0.2857
	Rank 4	0.5714	0.7143	0.8571	1.0000	0.8571	0.7143	0.5714	0.4286
	Rank 5	0.4286	0.5714	0.7143	0.8571	1.0000	0.8571	0.7143	0.5714
	Rank 6	0.2857	0.4286	0.5714	0.7143	0.8571	1.0000	0.8571	0.7143
	Rank 7	0.1429	0.2857	0.4286	0.5714	0.7143	0.8571	1.0000	0.8571
	Rank 8	0.0000	0.1429	0.2857	0.4286	0.5714	0.7143	0.8571	1.0000

The weighted observed proportional agreement *P(o)* can be expressed as Function 9.8: (Steltner et al., 2002)

$$P(o) = \frac{1}{N} \sum_{i=1}^{8} \sum_{j=1}^{8} w_{ij} n_{ij} \tag{9.2}$$

Let r_i be the sum of frequencies from Moody's rating and c_i the sum of frequencies from the present study's credit-scoring model rating. The weighted expected proportional agreement *P(e)* is estimated by:

$$P(e) = \frac{1}{N^2} \sum_{i=1}^{8} \sum_{j=1}^{8} w_{ij} r_i c_i \tag{9.3}$$

Finally, weighted Kappa is calculated with the following function:

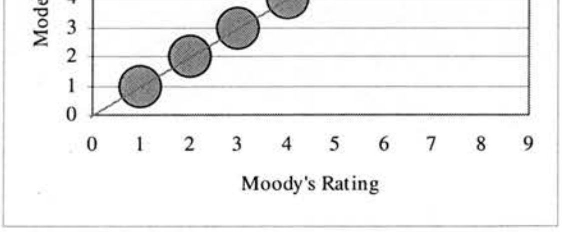
$$Kappa\ (w) = \frac{P(o) - P(e)}{1 - P(e)} \tag{9.4}$$

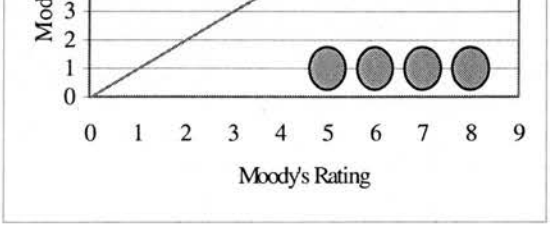
If the value of *Kappa (w)* is one, then the rating from the credit scoring model concords perfectly with Moody's rating. On the other hand, if the value of *Kappa (w)* is zero, then there is no concordance between the credit scoring model's rating and the Moody's rating. A negative value of *Kappa (w)* implies that the similarity is worse than the chance agreement (Steltner et al., 2002).

9.4.4 Graphical Bubble Charts

In this research, graphical analysis using the bubble chart was developed to facilitate the interpretation of similarity. The bubble chart enables a visualization of crosstabulation tables with clear localization of frequencies and a graphical representation of the observations through bubble size.

Bubble charts are interpreted as follows: The closer the bubbles are to the diagonal line, the more similar the rankings are. If the bubbles are close to the diagonal line are large in size, then it can be concluded that the degree of similarity between rankings is higher. If the bubbles are gathered in the upper left hand corner and in the lower right hand corner, then the degree of similarity between the compared rankings is low. Figures 9.1 and 9.2 are the examples of a best matching scenario and a worst matching scenario between compared rankings.





The above section introduced the credit scoring models for practical applicability evaluation, Moody's data collection and techniques for rating comparative analysis. The next two sections will compare Moody's ratings with the ratings from the credit scoring models based on PCA and Forward Stepwise variable selection approaches.

9.5 Ranking Comparison for PCA Models

9.5.1 Kolmogorov-Smirnov Test

The K-S test results are presented in Table 9.4:

Table 9.4 K-S Test Results for PCA Models

Methodology	K-S	2002	2001	2000	1999	1998
Logistic Regression	Z Value	0.917	1.241	1.225	1.128	1.490
	p-value	0.370	0.092	1.000	0.157	0.024
Neural Network	Z Value	1.250	1.067	2.123	4.271	2.979
	p-value	0.088	0.197	0	0	0
SMO	Z Value	2.583	1.490	1.061	3.062	1.821
	p-value	0	0.024	0.210	0	0.003

The highlighted *p*-values in Table 9.4 are not significant at a 5% level of significance and thus indicate when a proposed model provides rankings similar to Moody's. For example, Logistic Regression model has similar rankings in years 1999, 2000, 2001 and 2002. Neural Network model has similar rankings in years 2001 and 2002. SMO model has similar rankings in 2000. Therefore, based on the K-S test, it can be concluded that the Logistic Regression model displayed better performance than the other credit scoring models.

9.5.2 Distance Analysis

Distance analysis results are presented in Table 9.5:

Table 9.5 Distance Analysis Results for PCA Models

Methodology	2002	2001	2000	1999	1998	Average
Logistic Regression	1.5000	1.6986	1.6533	1.6623	1.6301	***1.6289***
Neural Network	1.9028	1.6164	1.8400	4.0130	3.7123	***2.6169***
SMO	1.6528	1.3836	1.3733	2.8672	1.7397	***1.8033***

Amongst the three models, the Neural Network model has the highest *average* distance, and the highest distances in the years 1998, 1999, 2000 and 2002. The Logistic Regression model performed best, with the lowest average distance over the five years (1.6289). However, the same cannot be concluded in each specific year. For example, SMO model displays lowest distance (and therefore the best performance) in 2000 and 2001. In fact, it can be concluded that the SMO model has only slightly worse performance than the Logistic Regression model, despite a higher average distance.

9.5.3 Weighted Kappa

Weighted Kappa results are presented in Table 9.6:

Table 9.6 Weighted Kappa Analysis Results for PCA Models

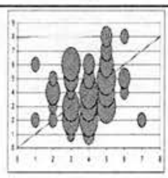
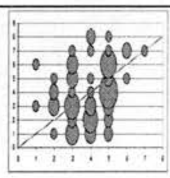
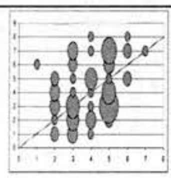
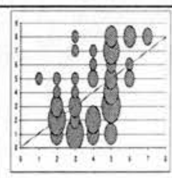
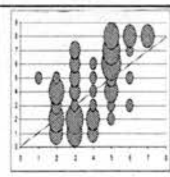
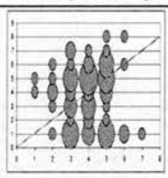
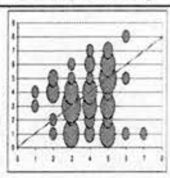
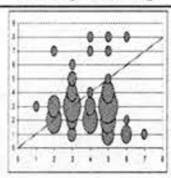
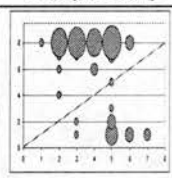
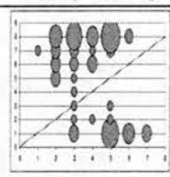
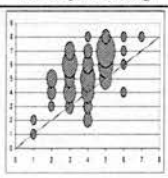
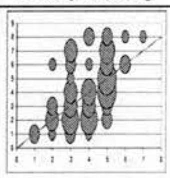
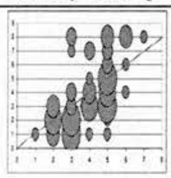
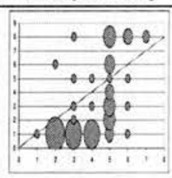
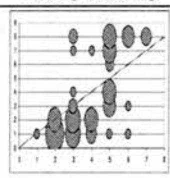
Methodology	2002	2001	2000	1999	1998	Average
Logistic Regression	0.1019	0.1073	0.1413	0.2664	0.2939	***0.1822***
Neural Network	0.0098	0.0980	0.0182	-0.1416	-0.1921	***-0.0415***
SMO	0.1762	0.3078	0.3496	0.0869	0.3403	***0.2522***

As with distance analysis results, Neural Network model still shows the worst performance. This is in terms of both the average value and yearly values. Interestingly, the average weighted Kappa and the weighted Kappa in 1998 and 1999 of Neural Network model are negative. This implies that the similarity between the rating of Neural Network model and the Moody's rating is worse than the chance agreement. However, unlike in distance analysis, SMO model performs better than the Logistic Regression model here. SMO model displays the highest weighted Kappa in the years 1998, 2000, 2001 and 2002. In sum, based on the weighted Kappa analysis, SMO model presents the best performance.

9.5.4 Graphical Bubble Charts Analysis

The bubble chart analysis is a quick way of comparing the degree of similarity between different ranking methods. The bubble charts based on different time scales are presented in Figure 9.3:

Figure 9.3 Bubble Charts for PCA Models

Methodology	2002	2001	2000	1999	1998
Logistic Regression					
	Moody's Rating	Moody's Rating	Moody's Rating	Moody's Rating	Moody's Rating
Neural Network					
	Moody's Rating	Moody's Rating	Moody's Rating	Moody's Rating	Moody's Rating
SMO					
	Moody's Rating	Moody's Rating	Moody's Rating	Moody's Rating	Moody's Rating

From Figure 9.3, it is clear that the Neural Network model performs worse in 1998 and 1999 than in 2000, 2001, or 2002. 1998 and 1999 show a higher number of large bubbles away from the diagonal line. These findings confirm those obtained through distance and weighted Kappa analyses, where the Neural Network model displayed the worst performance in the 1998 and 1999. The same conclusions can be made for the SMO model in 1999. This shows that bubble charts can provide some basic insights about the similarity between two ordinal datasets.

To summarize, it can be concluded that out of the PCA models used, Logistic Regression model displayed better performance than other two credit scoring models in terms of the K-S test and the distance analysis. However, in weighted Kappa analysis, SMO excelled over the other models in most of the time scales. Neural Network showed the worst similarity with Moody's rating based on both the distance and the weighted Kappa analyses. This was also detected from the bubble chart presentations.

9.6 Ranking Comparison for Forward Stepwise Models

9.6.1 Kolmogorov-Smirnov Test

K-S test results for forward stepwise models are presented in Table 9.7:

Table 9.7 K-S Test Results for Forward Stepwise Models

Methodology	K-S	2002	2001	2000	1999	1998
Logistic Regression	Z Value	-0.116	-1.899	-0.364	-1.071	-0.121
	p-value	0.908	0.058	0.716	0.284	0.904
Neural Network	Z Value	2.583	2.897	2.041	1.934	1.903
	p-value	0	0	0	0.001	0.001
SMO	Z Value	1.083	1.407	1.551	1.289	1.324
	p-value	0.191	0.038	0.016	0.072	0.060

The ratings between the Moody's and the Logistic Regression model are similar in all time scales. Regarding the SMO model, the results indicated that it has similar ratings with Moody's in years 1998, 1999 and 2002. In contrast, the Neural Network model does not show any statistical similarity with Moody's in any time scale, and hence displays the worst performance in terms of the K-S test.

9.6.2 Distance Analysis

The distance analysis results for stepwise models are presented in Table 9.8:

Table 9.8 Distance Analysis Results for Forward Stepwise Models

Methodology	2002	2001	2000	1999	1998	Average
Logistic Regression	1.0972	1.3288	1.3467	1.4416	1.3425	***1.3114***
Neural Network	1.5694	1.7397	1.6133	1.5844	1.3699	***1.5753***
SMO	1.0278	1.3014	1.3867	1.3896	1.3288	***1.2869***

Amongst the three models, the Neural Network model has the highest average distance between 1998 and 2002 as well as the highest distances each year. The Neural Network model therefore shows the worst performance. The best model is the SMO model based on average distance over the five years. The Logistic Regression model has similar performance to SMO model, although the average distance is slightly higher.

9.6.3 Weighted Kappa Analysis

The weighted Kappa results are expressed in Table 9.9:

Table 9.9 Weighted Kappa Analysis Results for Forward Stepwise Models

Methodology	2002	2001	2000	1999	1998	Average
Logistic Regression	0.4135	0.3338	0.3676	0.3529	0.4106	***0.3757***
Neural Network	0.2499	0.2164	0.2874	0.3553	0.4264	***0.3071***
SMO	0.4262	0.3364	0.3575	0.3691	0.4255	***0.3829***

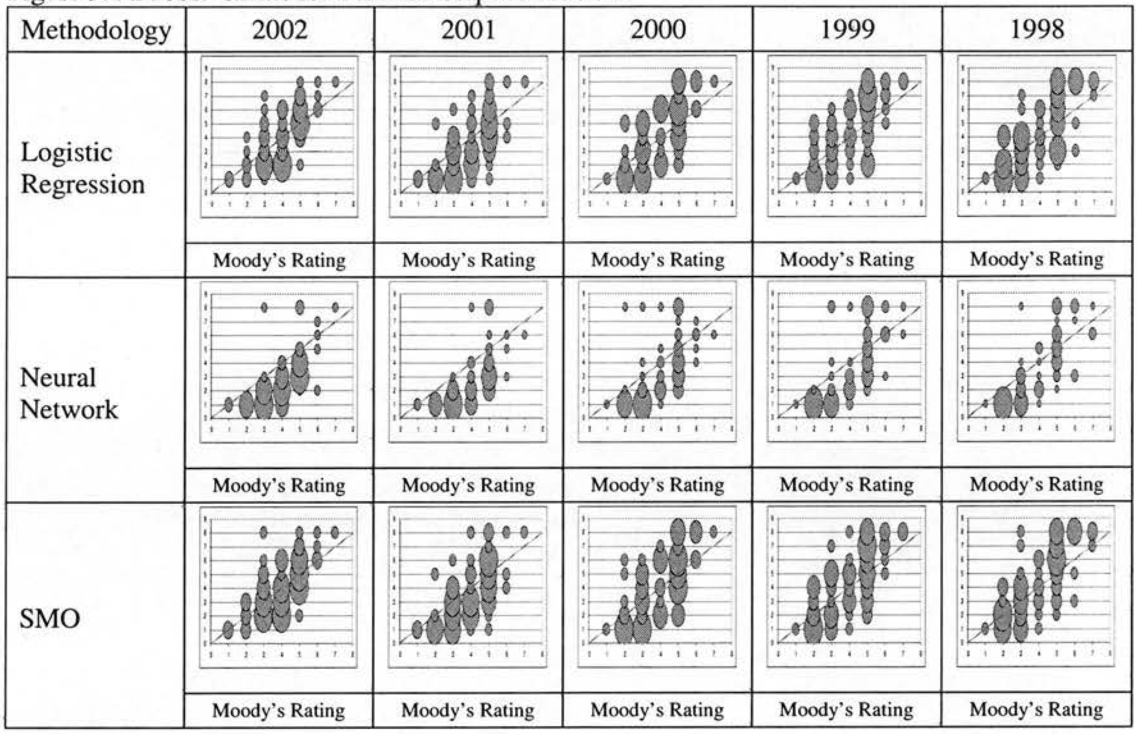
As with distance analysis, average weighted Kappa results suggest that the SMO model shows the highest degree of agreement with Moody's. The Neural network model still shows the lowest level of agreement with Moody's.

9.6.4 Graphical Bubble Charts Analysis

The bubble charts for the forward stepwise models are presented in the Figure 9.4. The bubbles distribution of the Neural Network model clearly stands out. The bubbles tend to locate below the diagonal line. Having bubbles above the diagonal line would have indicated Moody's provided better ratings than the research model. This would have meant that the research model underrates the credit situation for firms. In this case, having the bubbles below the diagonal line implies that the research model provide better ratings than Moody's. As a result, it can be concluded that the Neural Network model possibly overrates the credit situation of sample companies from 1998 to 2002. In addition, the Logistic Regression model and the

SMO model seem to have a very similar distribution of bubbles. This reflects findings from the distance and weighted Kappa analyses (These two credit scoring models show very similar performance).

Figure 9.4 Bubble Charts for Forward Stepwise Models



Although the bubble charts indicate more or less similarity between the research model's rating and Moody's rating, there are still a number of bubbles away from the diagonal line in Figure 9.3 and 9.4. Moody's rating methodology is based on the *'Through-the-Cycle'* perspective (Moody's Investor Service, 2002). Through-the-Cycle perspective has two implications: the ignorance of short-term fluctuations in default risk and a prudent rating migration policy. In terms of the first implication, Moody's rating approach focuses on the permanent, long-term and structural credit risk component. Therefore, the short-term changes in credit risk will not affect the rating immediately. The second implication means that only substantial changes in the permanent credit risk component will change the rating. Drawing on the above, the Through-the-Cycle perspective is characterized by a more stable and slow changing rating system, which may sometimes not be timely enough for investors and lenders when short-term changes in the market place arise.

Unlike Moody's approach, the *'Point-in-Time'* perspective was adapted in this research. This perspective, more useful to investors and lenders, concentrates on short-term credit risk and takes into account temporary credit risk components. The relationship between the Through-the-Cycle perspective and Point-in-Time perspective can be interpreted through bubble chart in this research. In Figure 9.4, although the bubble charts indicate more or less similarity between the research model's rating and Moody's rating, there are still a number of bubbles away from the diagonal line. The result indicates differences between the Through-the-Cycle perspective and Point-in-Time perspective. Looking at the bubble charts in more detailed, it can be noted that Moody's rating is more stable than research model's rating over time. For example, in terms of the 10th sample company, Moody's rating is fixed as 'A' from 1998 to 2002, while the ratings from logistic regression model changes from Baa, A, Baa, Baa, A in the same time period.

Another interesting finding is that unlike results in the previous chapter where the Logistic Regression and Neural Network models showed slightly better classification power than the SMO model based on average accuracy rate, SMO model's ability to rank company performance is slightly better than the Logistic Regression model and relatively better than the Neural Network model here. This is true for distance analysis and weighted Kappa analysis based on the Forward Stepwise approach as well as weighted Kappa analysis based on the PCA approach. An explanation is that the Neural Network and Logistic Regression model possibly overfit the sample, whilst the SMO model has not done so.

Thus far, this chapter has discussed the practical applicability of both PCA and Forward Stepwise default prediction models. When analyzing accuracy rate and AUROC value in the previous chapter, it was found that the Forward Stepwise models have better classification performance than PCA models (especially in the time periods one, two or three years before default). Can the same conclusion on practical applicability be drawn here, too? The following section will focus on this issue.

9.7 Comparative Analysis Between PCA and Stepwise Regression Models

9.7.1 Distance Comparative Analysis

The following line charts can be used to do the distance comparative analysis:

Figure 9.5 Logistic Regression Comparative Analysis (Distance)

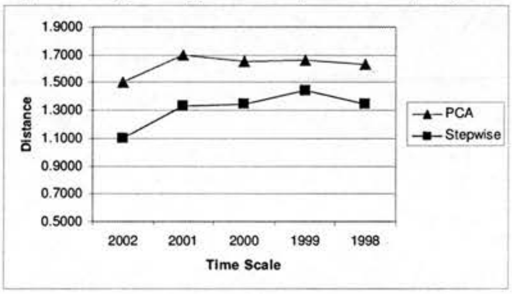


Figure 9.6 Neural Network Comparative Analysis (Distance)

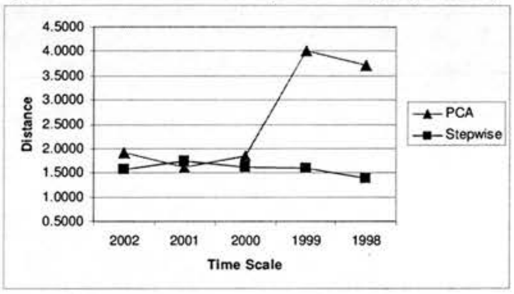
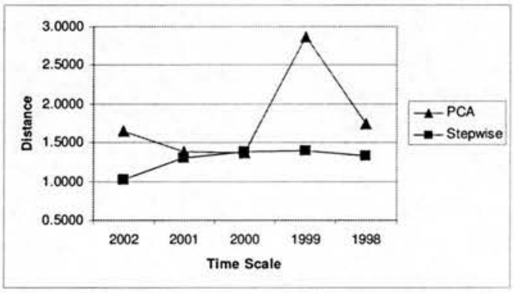


Figure 9.7 SMO Comparative Analysis (Distance)



Figures 9.5 to 9.7 show that most PCA models have higher distance than Forward Stepwise models. The only exceptions are Neural Network model in 2001 and SMO model in 2000. Average distance comparison shows the same: all Forward Stepwise models have lower average distance values than the PCA models. Furthermore, Forward Stepwise models show a more stable performance than PCA models, as the Neural Network PCA model and the SMO PCA model display a spike in 1999 (see Figure 9.7). Drawing on the analysis, it can be concluded that the Forward Stepwise models have better practical applicability than the PCA models.

9.7.2 Weighted Kappa Comparative Analysis

Line charts are again used to compare PCA and Forward Stepwise approaches using weighted Kappa.

Figure 9.8 Logistic Regression Comparative Analysis (Weighted Kappa)

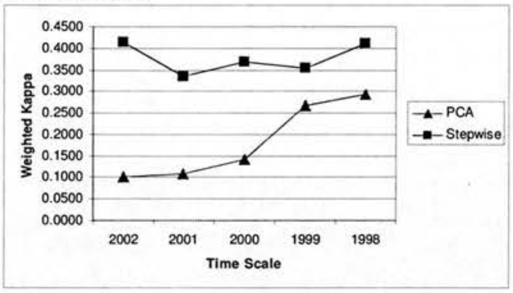


Figure 9.9 Neural Network Comparative Analysis (Weighted Kappa)

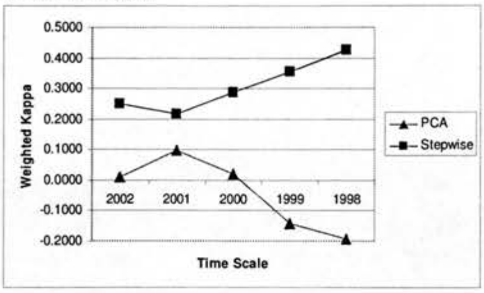
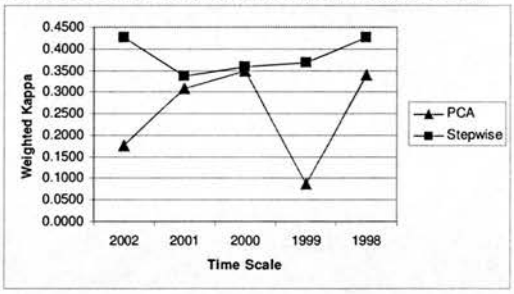


Figure 9.10 SMO Comparative Analysis (Weighted Kappa)



From the figures, higher weighted Kappa is seen in all the Forward Stepwise models. The performance of the Forward Stepwise model is also more stable than the PCA model in weighted Kappa analysis. Therefore, in terms of the weighted Kappa comparative analysis, the Forward Stepwise still outperforms than the PCA models.

Based on the findings in this chapter and the previous one, it can be concluded that the Forward Stepwise models have better performance than the PCA models in terms of both model prediction ability as well as of model practical applicability. The finding illustrates that the original research idea to use Moody's rating as a benchmark for measuring a default prediction model's practical applicability seems reasonable, since a consistent conclusion was obtained based on both the model prediction power and comparison with Moody's.

9.8 Concluding Remarks

This chapter evaluated the default prediction model's practical applicability through comparison with Moody's rating. Logistic Regression, Neural Network and

SMO models in the time period one year before financial distress in the variable group G12 were selected for the analysis. As this research employed two variable selection approaches, six credit scoring models were compared with Moody's. Four techniques: Kolmogorov-Smirnov (K-S) test, Distance analysis, Weighted Kappa analysis and Graphical Bubble charts were employed for the comparative analysis.

Regarding PCA models, Neural Network model showed the least similarity with Moody's rating in both distance and weighted Kappa analyses as well as in the bubble charts. Logistic Regression model displayed better performance than other two credit scoring models in terms of the K-S test and the distance analysis. However, SMO presents the best performance in terms of the average weighted Kappa and in most of the time scales.

With regards to Forward Stepwise models, SMO model's ability to rank company performance was slightly better than Logistic Regression and much better than Neural Network in distance analysis and weighted Kappa measure of agreement, but not in the K-S test. In the latter, the ratings between the Moody's and the Logistic Regression model are similar in all time scales. The bubbles distribution also presented similar results.

The findings above show a paradoxical result. On the one hand, the Logistic Regression and Neural Network models show slightly better model prediction power than the SMO model based on average accuracy rate. On the other hand, as discussed at length in this chapter, SMO model's ability to rank company performance is appeared to be better than both Logistic Regression model (to a small degree) and the Neural Network model (to a large degree) based on distance analysis and weighted Kappa analysis. An explanation is that the Neural Network model and the Logistic Regression model fit closely to, or possibly overfit the sample but the SMO model did not.

Finally, a comparative analysis between PCA and Stepwise Regression models using weighted Kappa and distance analyses gave a similar result as the model prediction ability assessment in the Chapter Eight: Forward Stepwise models outperform PCA models. The results pointed out that the original research idea to compare with Moody's rating with the aim of evaluating a default prediction model's practical applicability is reasonable. The conclusions also suggested that the five key variables: *Debt Ratio*, *Total Debt / (Total Debt + Market Capitalization)*, *Total Assets*, *Operating Cash Flow* and *Government Debt / GDP* have better classification ability and practical applicability than the principal components.

As mentioned in the Chapter One, the model's practical applicability can also be assessed by applying the original US model to different markets. Another interesting research question is: *'Are these five variables still useful in different retail markets?'* European and Japanese markets are selected to answer this question. This allows for international comparison of the model's prediction performance in different contexts. Such will be the subject in the next chapter.

Chapter TEN

Model Practical Applicability Evaluation: International Applicability

10.1 Introduction

In the previous two chapters, the focus was on comparing PCA and Stepwise Regression models. Results indicated that the Forward Stepwise models have better prediction performance than the PCA models. In other words, the five key variables: *Debt Ratio*, *Total Debt / (Total Debt + Market Capitalization)*, *Total Assets*, *Operating Cash Flow* and *Government Debt / GDP* have better prediction performance than the principal components. The next issue to explore is whether the five variables have good prediction power in different retail markets. European and Japanese market data were used to do this. Moreover, the model was applied to a new US data set for comparison purposes.

This chapter starts by introducing the new markets' data and the sample selection criteria. This is followed by an analysis of prediction power in each single market. A comparative analysis among different markets is then presented for the purpose of comparing cross-border performance. A summary is found in the final section.

10.2 Data Collection

10.2.1 Sample Selection Criteria

With regards to the sample selection of *healthy* firms, three criteria were considered. First, only listed firms were included. Listed companies need to obey the regulations in the financial market and therefore provide more transparent data. Another decision on sample was to omit e-retailers, because the performance measures of e-retailers are different. Finally, even if a company satisfied the two criteria above, it would still be excluded if its data is not complete.

Selection of *distressed* companies is based on financial criteria. Ross, et al. (1999) pointed out the definition of financial distress has two themes: stock-based insolvency and flow-based insolvency. Stock-based insolvency occurs when a company's total liabilities are greater than its total assets. Flow-based insolvency occurs when a company's operating cash flow cannot meet its routine obligations. Hence, a company was regarded as distressed in this research when its debt to equity ratio was negative (stock-based insolvency) or when its interest cover based on cash flow framework (EBITDAR / interest) was smaller than one (flow-based insolvency).

10.2.2 Data Description

Thomson One Banker database was the main data source for company financial data. Macroeconomical data was collected from Organisation for Economic Co-operation and Development (OECD) documents. Table 10.1 summarises the data of the three target markets from 2000 to 2004. (The sample composition in the Table 10.1 will be labelled as the '*Original Data Set*' in this chapter.)

Table 10.1 Original Data Description

Target markets	2004		2003		2002		2001		2000	
	Healthy	Distressed	Healthy	Distressed	Healthy	Distressed	Healthy	Distressed	Healthy	Distressed
USA	181	24	179	40	190	46	184	63	190	70
Europe[1]	145	27	162	26	164	31	182	32	195	31
Japan	251	28	244	19	219	17	180	55	195	39
Total	577	79	585	85	573	94	546	150	580	140

An initial interest of this study was the time scale effect—whether one should use data just before the default or some time before. Setting 2004 as the year prior to financial distress could allow for time scale effect detection. For example, 2003 may then regard as the time period two years before financial distress, 2002 as the time period three years before financial distress, and so on. As a result, only companies with five years complete data were considered when exploring time scale effects. The sample size of each country is illustrated in Table 10.2:

1 The composition of the European market includes the 25 countries in the European Union plus Swaziland and Norway.

Table 10.2 Data Description for Exploring Time Scale Effect

	USA	European	Japan	Total
Healthy	170	126	195	491
Distressed	21	20	27	68

Again, the Classification Accuracy Rate and the Area under the Receiver Operating Characteristics Curve (AUROC) were employed to assess the classification ability of each market. The performance of each market will be discussed in turn.

10.3 Evaluation of Classification Power for US New Model

10.3.1 Accuracy Rate Analysis

10.3.1.1 Original Data Comparative Analysis

Based on the original data in Table 10.1, the accuracy rates of each credit scoring technique from 2000 to 2004 are shown in the Table 10.3:

Table 10.3 Original Data Accuracy Rate Comparative Analysis (US New Model)

Methodology	2004	2003	2002	2001	2000	Average
Naïve Bayes	89.76%	90.41%	91.10%	87.85%	86.54%	***89.13%***
Logistic Regression	92.20%	90.87%	90.25%	87.85%	87.69%	***89.77%***
Neural Network	91.22%	87.21%	88.14%	87.04%	85.77%	***87.88%***
SMO	92.68%	89.95%	89.83%	87.04%	82.31%	***88.36%***
Recursive Partitioning	93.17%	88.58%	88.56%	82.59%	83.85%	***87.35%***

The higher the accuracy rate is, the better the performance. Different credit scoring techniques display different performance in terms of different time scales. For example, Recursive Partioning model shows the best performance in 2004 but not in the other years. By comparing the average performance among five credit scoring models, the Logistic Regression model gives the highest average accuracy rate of 89.77%. Nevertheless, the difference among five credit scoring models is very small (below 3%). Therefore, it should be said that the five credit scoring models present similar performance based on the US new data. Given that the average accuracy rate is above 87.35%, regardless of the credit scoring technique employed, it can also be argued that the five key variables, *Debt Ratio*, *Total Debt / (Total Debt + Market Capitalization)*, *Total Assets*, *Operating Cash Flow* and *Government Debt / GDP*, show sound prediction power.

10.3.1.2 Exploring Time Scale

Based on the time scale data in Table 10.2, let 2004 be the year before financial distress, M1, 2003 two years prior to financial distress, M2, and 2002 three years before default, M3, and so on until M5. Results of the accuracy rate analysis within a five-year time scale are presented in Table 10.4.

Table 10.4 Exploring Time Scale: Accuracy Rate (US New Model)

Methodology	Performance	M1	M2	M3	M4	M5	Average
Naïve Bayes	Type I Error	33.33%	38.10%	52.38%	33.33%	28.57%	***37.14%***
	Type II Error	6.47%	5.88%	2.94%	7.65%	10.00%	***6.59%***
	Overall	90.58%	90.58%	91.62%	89.53%	87.96%	***90.05%***
Logistic Regression	Type I Error	47.62%	52.38%	57.14%	42.86%	47.62%	***49.52%***
	Type II Error	2.94%	4.12%	2.35%	2.94%	2.35%	***2.94%***
	Overall	92.15%	90.58%	91.62%	92.67%	92.67%	***91.94%***
Neural Network	Type I Error	61.90%	66.67%	52.38%	47.62%	52.38%	***56.19%***
	Type II Error	3.53%	5.88%	3.53%	2.94%	2.35%	***3.65%***
	Overall	90.05%	87.43%	91.10%	92.15%	92.15%	***90.58%***
SMO	Type I Error	71.43%	80.95%	80.95%	42.86%	61.90%	***67.62%***
	Type II Error	0.59%	0.00%	0.00%	2.94%	1.76%	***1.06%***
	Overall	91.62%	91.10%	91.10%	92.67%	91.62%	***91.62%***
Recursive Partitioning	Type I Error	52.38%	61.90%	61.90%	66.67%	66.67%	***61.90%***
	Type II Error	2.35%	0.59%	3.53%	2.35%	1.18%	***2.00%***
	Overall	92.75%	92.67%	90.05%	90.58%	91.62%	***91.53%***

Regardless of the credit scoring technique employed, the accuracy rates are reasonably high: above 90.05% one year before financial distress (M1) and above

87.96% five years prior to default (M5). All credit scoring models display a very similar performance based on each specific year, but also based on the five years average performance. For example, the difference in average performance among different credit scoring techniques is below 2%. Therefore, the same conclusion can be made as in Chapter Eight—that it is difficult to conclude which credit scoring technique has the *'absolute'* best performance, since the performance varies in terms of different time scale and the differences among five modelling techniques are small.

10.3.1.3 Types of Error

In relation to types of error, Table 10.4 shows that the SMO model *'still'* has the best ability to deal with the Type II error and the worst with Type I error. The Naïve Bayes model displays the lowest value of Type I error not only based on each specific year, but also based on the average performance.

10.3.2 AUROC Analysis

10.3.2.1 Original Data Comparative Analysis

The AUROC values based on the original data are expressed in Table 10.5:

Table 10.5 Original Data AUROC Comparative Analysis (US New Model)

Methodology	2004	2003	2002	2001	2000	Average
Naïve Bayes	0.9332	0.9345	0.9418	0.9180	0.9238	***0.9303***
Logistic Regression	0.9399	0.9426	0.9141	0.9137	0.9168	***0.9254***
Neural Network	0.8946	0.8997	0.8715	0.8992	0.8944	***0.8919***

Table 10.5 displays the worst classification performance for the Neural Network model based on yearly value or average value. Between Naïve Bayes model and Logistic Regression, the average performance of Naïve Bayes model is better than Logistic Regression model, although the difference is very small. Moreover, regardless of the credit scoring technique employed, the average performance is

above 0.8919. This implies that the five variables still have sound prediction performance when using the new USA data set. The conclusion to be drawn is, the AUROC performance of Naïve Bayes model is almost the same as Logistic Regression model and slightly better than Neural Network model.

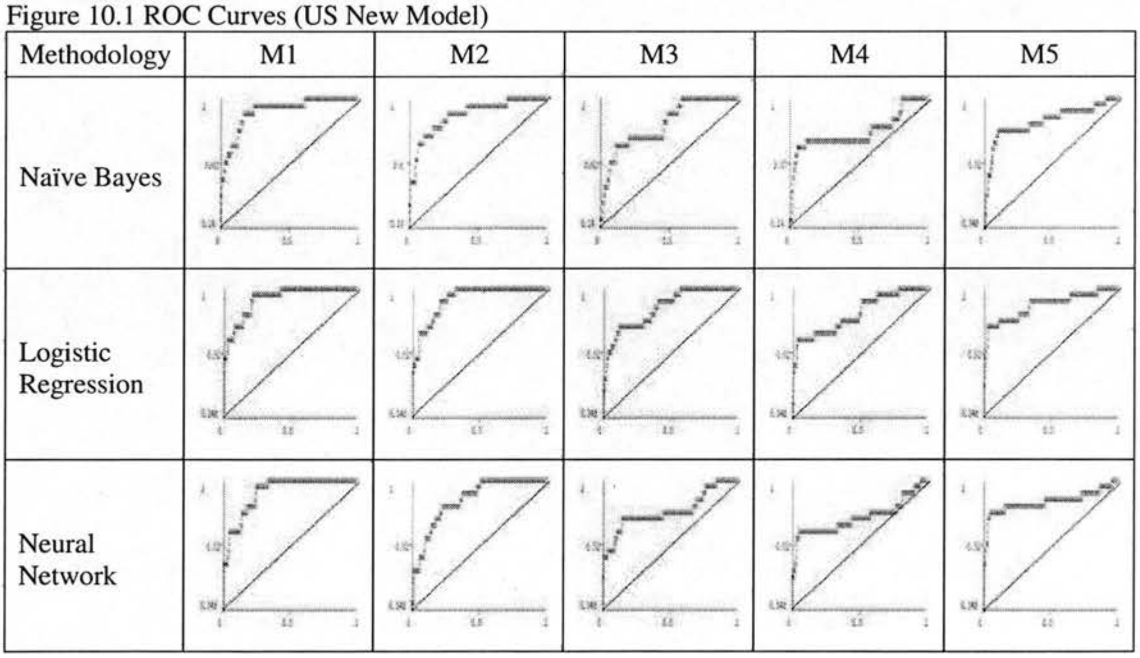
10.3.2.2 Exploring Time Scale

The AUROC values and the ROC curves based on time scale data are presented in Table 10.6 and Figure 10.1:

Table 10.6 Exploring Time Scale: AUROC (US New Model)

Methodology	M1	M2	M3	M4	M5	Average
Naïve Bayes	0.9238	0.8964	0.8454	0.7751	0.8210	***0.8523***
Logistic Regression	0.9241	0.9210	0.8555	0.8123	0.8709	***0.8768***
Neural Network	0.9087	0.8524	0.7714	0.7218	0.8339	***0.8176***

Figure 10.1 ROC Curves (US New Model)



From Table 10.6 and Figure 10.1, it is clear that all the credit-scoring models show the best performance one year before financial distress with an AUROC value of above 0.9087. Moreover, the Logistic Regression model presents the highest *average*

AUROC value. Despite superior performance of the Logistic Regression model, the difference among three credit scoring models is small.

Another interesting discovery is that the worst performance of each credit scoring model is in M4, not M5. This is also be detected through the ROC curves. In addition, notwithstanding the credit scoring techniques employed, the AUROC value in M4 is still above 0.7218. Based on the definition in Hosmer and Lemeshow (2000), such discriminating power is still acceptable. Looking at average performance, it can be seen that the average AUROC is above 0.8176 despite modelling techniques. The above provides sufficient evidences that the five key variables have sound classification ability in the US market based on a new data set.

10.3.3 Concluding Remarks for US New Model Analysis

In US original new dataset, the Logistic Regression model showed the highest average accuracy rate (89.77%) and the Naïve Bayes model displayed the best average AUROC value (0.9303). However, the difference among credit scoring models is very small. (The difference in accuracy rates is below 0.3% and in AUROC values is below 0.04). Therefore, it can be concluded that all five techniques have similar performance based on the US original dataset. Furthermore, the results also illustrate that the five variables have sound prediction power.

On the topic of time scale, in spite of the performance measures, all credit scoring models display very similar performance not only based on each specific year, but also based on five years average performance. From the accuracy rate analysis, it is difficult to conclude which credit scoring technique has the *'absolute'* best performance, since the performance results vary according to time scale and the difference among the five credit scoring techniques remains small. However, the Logistic Regression model presents the best performance in terms of the AUROC value. Although the Logistic Regression model shows the highest AUROC values, there is little difference among the three credit scoring models. As the average accuracy rate is above 90.05% and the average AUROC value is above 0.8176 for the new USA dataset, there is sufficient evidence that the five key variables truly have sound classification ability.

An analysis on types of error was carried out. It was found that SMO model was the best model in dealing with the Type II error and the worst one for Type I error. Furthermore, the Naïve Bayes model displayed the best ability to cope with the Type I error in each specific year, and also on average. Such a conclusion is consistent with the results in Chapter Eight.

10.4 Evaluation of Classification Power for European Model

10.4.1 Accuracy Rate Analysis

10.4.1.1 Original Data Comparative Analysis

The accuracy rates of each credit scoring technique from 2000 to 2004 based on the original data can be expressed in Table 10.7:

Table 10.7 Original Data Accuracy Rate Comparative Analysis (European Model)

Methodology	2004	2003	2002	2001	2000	Average
Naïve Bayes	88.95%	89.89%	84.62%	87.85%	89.38%	***88.14%***
Logistic Regression	90.12%	88.30%	90.26%	90.19%	90.71%	***89.92%***
Neural Network	88.95%	87.77%	83.59%	90.65%	92.04%	***88.60%***
SMO	88.95%	86.70%	84.62%	85.51%	88.05%	***86.77%***
Recursive Partitioning	91.28%	89.36%	85.13%	87.38%	87.61%	***88.15%***

The results based on the European dataset in Table 10.7 show a conclusion similar to the US new dataset. The Logistic Regression model displays the best performance in terms of average accuracy rate, but the difference among five credit scoring techniques is small (below 4%). Moreover, regardless of the credit scoring technique used, the average accuracy rate is still above 86.77%. Hence, all five key variables have good prediction power in the European market.

10.4.1.2 Exploring Time Scale

The results of the accuracy rate based on a five-year time scale are presented in Table 10.8.

Table 10.8 Exploring Time Scale: Accuracy Rate (European Model)

Methodology	Performance	M1	M2	M3	M4	M5	Average
Naïve Bayes	Type I Error	40.00%	40.00%	60.00%	65.00%	90.00%	***59.00%***
	Type II Error	5.56%	7.14%	7.14%	4.76%	2.38%	***5.40%***
	Overall	89.73%	88.36%	85.62%	86.99%	85.62%	***87.26%***
Logistic Regression	Type I Error	60.00%	65.00%	90.00%	100.00%	100.00%	***83.00%***
	Type II Error	3.97%	3.17%	1.59%	0.00%	1.59%	***2.06%***
	Overall	88.36%	88.36%	86.30%	86.30%	84.93%	***86.85%***
Neural Network	Type I Error	50.00%	65.00%	65.00%	80.00%	95.00%	***71.00%***
	Type II Error	2.38%	3.17%	6.35%	1.59%	4.76%	***3.65%***
	Overall	91.10%	88.36%	85.62%	87.67%	82.88%	***87.13%***
SMO	Type I Error	65.00%	100.00%	100.00%	100.00%	100.00%	***93.00%***
	Type II Error	0.00%	0.00%	0.00%	0.00%	0.00%	***0.00%***
	Overall	91.10%	86.30%	86.30%	86.30%	86.30%	***87.26%***
Recursive Partitioning	Type I Error	55.00%	65.00%	75.00%	85.00%	100.00%	***76.00%***
	Type II Error	3.17%	3.97%	0.79%	0.79%	0.00%	***1.74%***
	Overall	89.73%	87.67%	89.04%	87.67%	86.30%	***88.08%***

From Table 10.8, all credit scoring models present the best performance in the year before financial distress. In addition, regardless of the modelling technique, accuracy rates are above 88.36% in M1. The accuracy rates still remain sound five years before financial distress (over 82.88%) for all modelling approaches. Recursive Partioning model presents the highest average accuracy rate (88.08%) but the differences with the other credit scoring techniques are small (below 2%). Furthermore, the performance fluctuates in terms of different time periods. Therefore, the same conclusion can be drawn again. Each credit-scoring technique has similar performance based on the European market and it is difficult to conclude which modelling method has the absolute best performance.

10.4.1.3 Types of Error

The analysis of Types of error for European market data has the same results as those for the US new market: the SMO model shows the best performance to deal with the Type II error but the worst for Type I error. Nevertheless, the results also indicate that the SMO model only has limited discriminating power based on Type I error from 2000 to 2003. It also implies that the SMO model tends to classify most sample companies into healthy firms. Again, the Naïve Bayes model still shows the best performance not only in terms of on the average accuracy rate but also based on each time scale.

10.4.2 AUROC Analysis

10.4.2.1 Original Data Comparative Analysis

The AUROC values based on the original European data (see Table 10.1) set are expressed in the Table 10.9:

Table 10.9 Original Data AUROC Comparative Analysis (European Model)

Methodology	2004	2003	2002	2001	2000	Average
Naïve Bayes	0.8835	0.8184	0.8387	0.9250	0.8996	***0.8730***
Logistic Regression	0.8733	0.8253	0.8324	0.9050	0.8951	***0.8662***
Neural Network	0.8248	0.8029	0.7710	0.8913	0.9179	***0.8416***

From Table 10.9, it is obvious that the three credit scoring techniques show different utility in different years. For example, Neural Network model presents the worst performance in 2004, but displays the best performance in 2000. On average, Naïve Bayes model shows the highest average AUROC value (0.8730). However, the differences of average AUROC value among modelling techniques are very small (below 0.04). As the average AUROC value is higher than 0.8416, the conclusion as the accuracy rate analysis can be made. The five variables still present good prediction power in the European market.

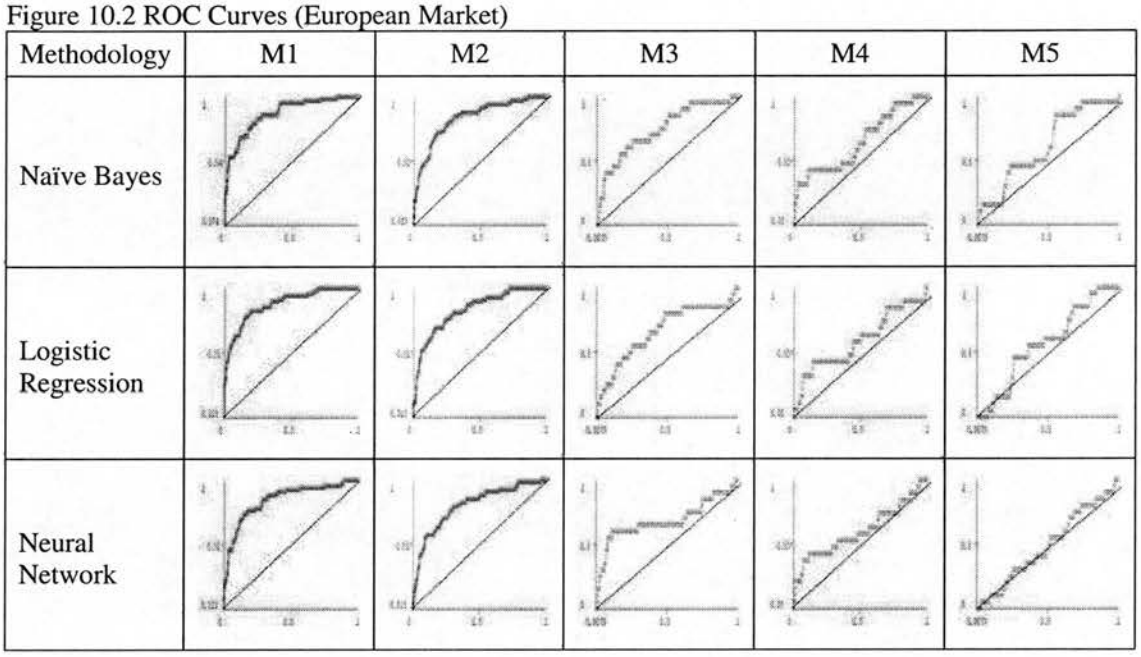
10.4.2.2 Exploring Time Scale

AUROC values and ROC curves based on the time scale data set are presented in Table 10.10 and Figure 10.2:

Table 10.10 Exploring Time Scale: AUROC (European Model)

Methodology	M1	M2	M3	M4	M5	Average
Naïve Bayes	0.8964	0.7813	0.7369	0.6603	0.6067	***0.7363***
Logistic Regression	0.8694	0.6619	0.6571	0.6052	0.5619	***0.6711***
Neural Network	0.8294	0.7298	0.6754	0.6159	0.5001	***0.6701***

Figure 10.2 ROC Curves (European Market)



As can be seen above, all credit-scoring models show the best performance in M1, the year before financial distress. However, unlike the results from the new US dataset, the prediction power is not satisfactory five years before financial distress. In M5, the AUROC value is only 0.6067 for Naïve Bayes model, 0.5619 for Logistic Regression model, and 0.5001 for Neural Network model in M5. Such is also illustrated in the ROC curves: in M5, the area between the ROC curve and the diagonal line is very small, in particular for Neural Network model. The above leads to the conclusion that the five variables have sound prediction power in the time period one year before financial distress, but the longer the time period prior to financial distress, the weaker the prediction power becomes, especially in the time period five years before financial distress.

10.4.3 Concluding Remarks for European Dataset Analysis

In analyzing the original European dataset, the research found an average accuracy rate of above 86.77% and an average AUROC value of above 0.8416. This implies that the five variables possess equally sound prediction power in the European market. The same can be concluded by viewing the average performance, since the average performance difference among the five credit scoring techniques is small

With regards to exploring time scale, it was discovered that all credit scoring models show the best performance one year before financial distress (in terms of either accuracy rate or AUROC value). However, in connection with the long term prediction power, there is a paradoxical result between the accuracy rate analysis and the AUROC analysis. The accuracy rate remains above 82.88% even if the time period is five years before default. In contrast, the long-term prediction power was not satisfactory based on AUROC value. For all credit scoring techniques, the AUROC values were below 0.6067 in M5. It was concluded that the five variables have sound prediction power in the time period one year before financial distress, but the longer the time period prior to financial distress, the weaker the prediction power becomes, particularly for the AUROC value.

On the subject of the types of error, although SMO still displays the best performance to deal with the Type II error and displays the worst performance to manage the Type I error, it only has limited classification ability based on the Type I error in most of the time periods. In other words, this implies that the five variables do not appear to have much classification utility in the European market based on the SMO approach. Regarding the ability to deal with the Type I error, Naïve Bayes model still shows the best performance not only based on the average accuracy rate but also in each year.

10.5 Evaluation of Classification Power for Japanese Model

10.5.1 Accuracy Rate Analysis

10.5.1.1 Original Data Comparative Analysis

The results of the accuracy rates based on the original Japanese dataset are shown in Table 10.11:

Table 10.11 Original Data Accuracy Rate Comparative Analysis (Japanese Model)

Methodology	2004	2003	2002	2001	2000	Average
Naïve Bayes	89.25%	91.25%	91.10%	79.57%	84.62%	***87.16%***
Logistic Regression	90.32%	91.64%	94.49%	80.00%	87.18%	***88.73%***
Neural Network	89.96%	92.02%	93.22%	78.30%	88.03%	***88.31%***
SMO	89.96%	92.78%	92.80%	78.72%	88.46%	***88.54%***
Recursive Partitioning	89.61%	93.16%	92.37%	76.60%	88.03%	***87.95%***

From Table 10.11, the Logistic Regression model presents the highest average accuracy rate, but the difference among modelling techniques is very small (below 2%) among the five credit scoring models. In addition, despite the credit scoring techniques, the average accuracy rate is above 87.16%. The findings still indicate that the five key variables have sound prediction power in the Japanese market.

10.5.1.2 Exploring Time Scale

The results of the accuracy rate based on a five-year time scale are presented in Table 10.12.

Table 10.12 Exploring Time Scale: Accuracy Rate (Japanese Model)

Methodology	Performance	M1	M2	M3	M4	M5	Average
Naïve Bayes	Type I Error	51.85%	70.37%	66.67%	62.96%	70.37%	***64.44%***
	Type II Error	5.64%	5.64%	6.15%	8.72%	7.18%	***6.67%***
	Overall	88.74%	86.49%	86.49%	84.68%	85.14%	***86.31%***
Logistic Regression	Type I Error	81.48%	96.30%	81.48%	96.30%	100.00%	***91.11%***
	Type II Error	2.56%	2.56%	2.05%	1.54%	1.03%	***1.95%***
	Overall	87.84%	86.04%	88.29%	86.94%	86.94%	***87.21%***
Neural Network	Type I Error	85.19%	92.59%	51.85%	88.89%	100.00%	***83.70%***
	Type II Error	3.08%	4.62%	5.13%	6.15%	1.03%	***4.00%***
	Overall	86.94%	84.68%	89.19%	83.78%	86.94%	***86.31%***
SMO	Type I Error	100.00%	100.00%	100.00%	100.00%	100.00%	***100.00%***
	Type II Error	0.00%	0.00%	0.00%	0.00%	0.00%	***0.00%***
	Overall	87.84%	87.84%	87.84%	87.84%	87.84%	***87.84%***
Recursive Partitioning	Type I Error	70.37%	100.00%	51.85%	88.89%	96.30%	***81.48%***
	Type II Error	7.18%	3.59%	6.67%	1.54%	1.54%	***4.10%***
	Overall	85.14%	84.68%	87.84%	87.84%	86.94%	***86.49%***

As illustrated in Table 10.12, accuracy rates remain similar in different time periods. For example, regardless of five modelling techniques, the lowest accuracy rate in M1 and M5 are identical: 85.14%. As a result, time scale effects are not obvious in the Japanese market. By comparing average performance, the difference is very small (below 2%) among different credit scoring models.

10.5.1.3 Types of Error

Similar to the conclusions for the European market analysis, the five variables applied to the Japanese dataset only have limited classification utility based on the SMO technique. Indeed, the SMO model does not display any discriminating ability to deal with Type I error in any period. Naïve Bayes model still shows the best average performance to manage the Type I error in most of the time periods. The only exception is M3, where the Neural Network and Recursive Partitioning models show better performance than Naïve Bayes.

10.5.2 AUROC Analysis

10.5.2.1 Original Data Comparative Analysis

The AUROC values for the Japanese dataset are expressed in the Table 10.13:

Table 10.13 Original Data AUROC Comparative Analysis (Japanese Model)

Methodology	2004	2003	2002	2001	2000	Average
Naïve Bayes	0.8552	0.8013	0.8020	0.8545	0.8368	***0.8300***
Logistic Regression	0.8314	0.8218	0.7814	0.8515	0.8393	***0.8251***
Neural Network	0.8059	0.7677	0.7806	0.8098	0.7888	***0.7906***

The Naïve Bayes model in Table 10.13 shows the best average performance among the three techniques, but the differences are small. The average AUROC value is above 0.7906 implying good classification ability for the five variables.

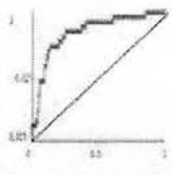
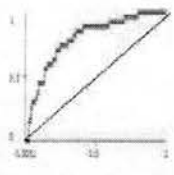
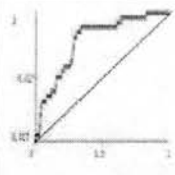
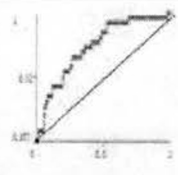
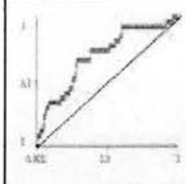
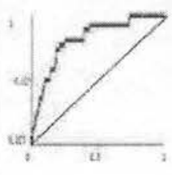
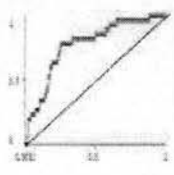
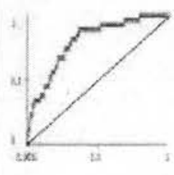
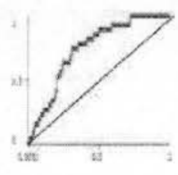
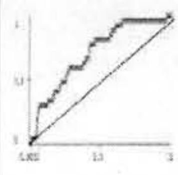
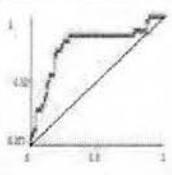
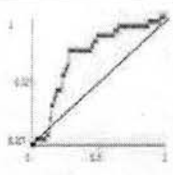
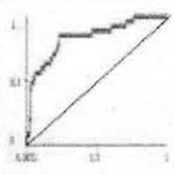
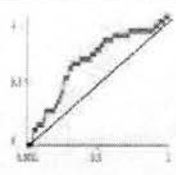
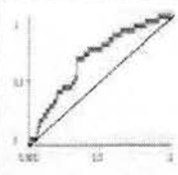
10.5.2.2 Exploring Time Scale

Results of AUROC values and ROC curves based on a five-year time scale are presented in Table 10.14 and Figure 10.3.

Table 10.14 Exploring Time Scale: AUROC Value (Japanese Model)

Methodology	M1	M2	M3	M4	M5	Average
Naïve Bayes	0.8454	0.7909	0.7837	0.7417	0.6948	***0.7713***
Logistic Regression	0.8184	0.7649	0.7928	0.7358	0.7005	***0.7625***
Neural Network	0.7725	0.6999	0.8342	0.6443	0.6615	***0.7225***

Figure 10.3 ROC Curves (Japanese Market)

Methodology	2004	2003	2002	2001	2000
Naïve Bayes					
Logistic Regression					
Neural Network					

Naïve Bayes and Logistic Regression models show the highest AUROC value in the year before financial distress (see Table 10.14). However, for Neural Network model, the best performance appears in M3. The results also indicate that the AUROC value in the time scale of five years before financial distress is above 0.6615 for all credit scoring techniques. Although the classification ability is not satisfactory, the result is still acceptable. With regards to the average performance, all credit-scoring models present a similar result although Naïve Bayes model shows slightly better performance. This finding is consistent with the US new model analysis and the European model analysis.

10.5.3 Concluding Remarks for Japanese Model Analysis

Using the original dataset, Logistic Regression model and Naïve Bayes model show the best performance based on the average accuracy rate and the average AUROC value respectively. At the same time, the differences among five credit scoring techniques are small. In other words, the performance among different credit scoring models is similar. The average accuracy rate is above 87.16% and the average AUROC value is above 0.7906 for all models. This means that the five key variables still possess sound prediction power in the Japanese market.

Time scale effects are not observable in the Japanese market based on accuracy rates. Indeed, accuracy rates do not vary much in different time scales nor by modelling technique. For AUROC analysis, not all the credit scoring models present the best performance in the year before financial distress. For example, Neural Network model displays the best performance three years prior to default with the AUROC value (0.8342). Finally, the AUROC value in the time scale of five years before financial distress is above 0.6615 in all cases. Although the result is not satisfactory, it is still acceptable.

Types of error analysis gave results different from those of the European market analysis. The five variables do not show classification utility based on the SMO technique, as the latter does not display any discriminating ability to manage Type I error. The Naïve Bayes model still shows the best performance to manage Type I error on average and in most time periods. Up till now, this chapter has discussed the prediction performance of different credit scoring models using US, European or Japanese market data. The next section will concentrate on a cross-border comparative analysis based on the three target markets.

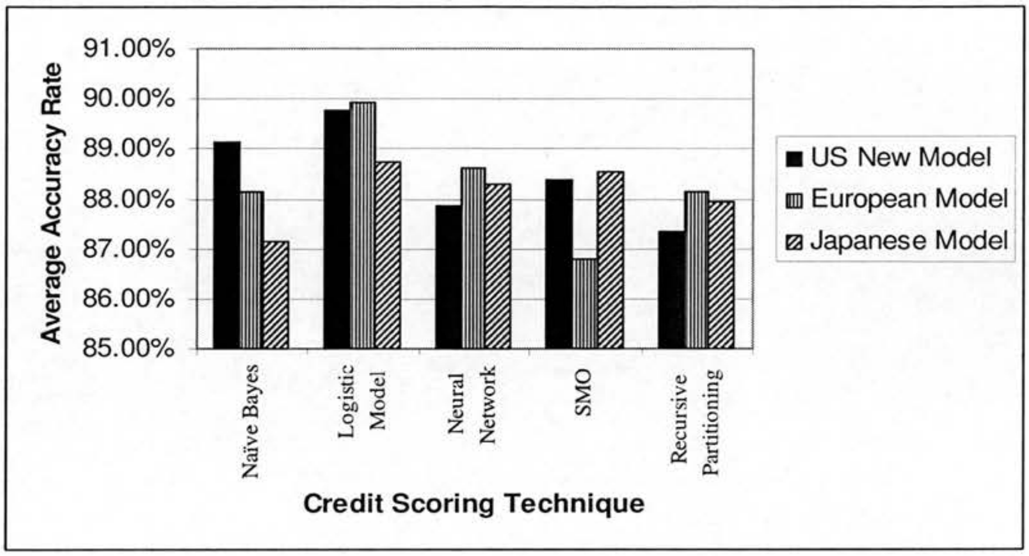
10.6 Cross-Border Comparative Analysis

10.6.1 Accuracy Rate Analysis

10.6.1.1 Original Data Comparative Analysis

The following bar chart is employed to conduct the comparative analysis based on the average accuracy rate:

Figure 10.4 Original Data Comparative Analysis (Average Accuracy Rate)



From Figure 10.4, it is difficult to conclude which market model has the absolute best average accuracy rate. For example, USA new model shows the best performance using the Naïve Bayes technique, but not in terms of all credit scoring techniques. In addition, the average accuracy rates range between 86.77% and 89.77%. This implies that the difference in performance in the three markets is very small. Therefore, it can be concluded that the three market models have similar average accuracy rate based on the original dataset.

10.6.1.2 Exploring Time Scale

The following line charts can be used to explore the time scale effects among different market models.

Figure 10.5 Naïve Bayes Comparative Analysis (Accuracy Rate)

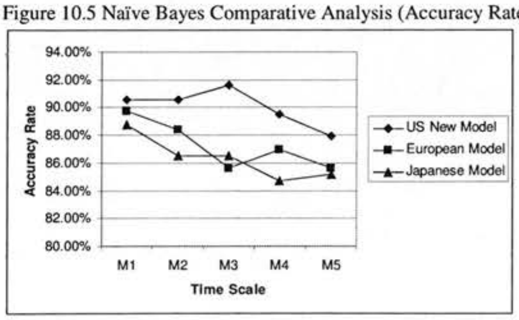


Figure 10.6 Logistic Regression Comparative Analysis (Accuracy Rate)

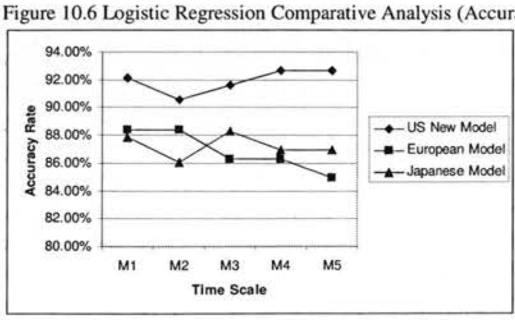


Figure 10.7 Neural Network Comparative Analysis (Accuracy Rate)

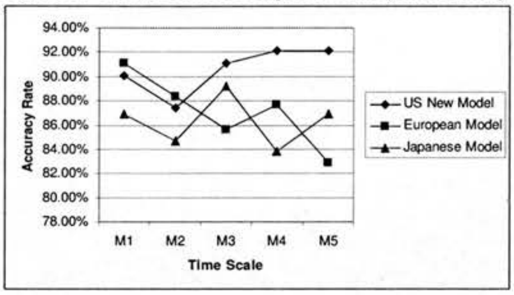


Figure 10.8 SMO Comparative Analysis (Accuracy Rate)

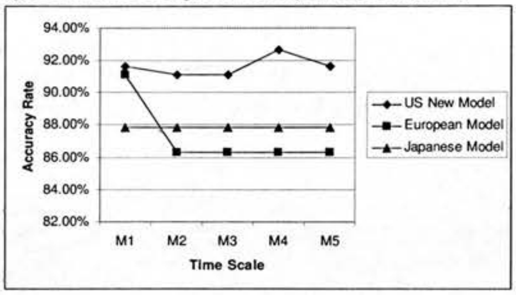
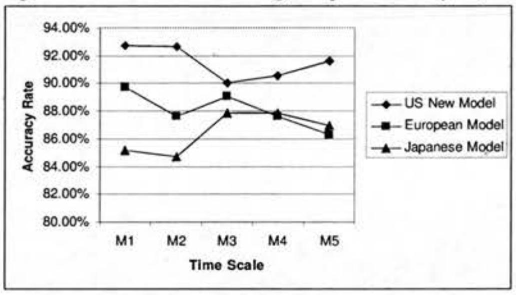


Figure 10.9 Recursive Partitioning Comparative Analysis (Accuracy Rate)



The new USA model has the best performance in almost all of the time periods. The only special case is the European Neural Network model, which displays the best performance in M1 and M2. However, between European and Japanese models, it is difficult to conclude which model is better because the model's performance changes depending on the modelling technique and time scale.

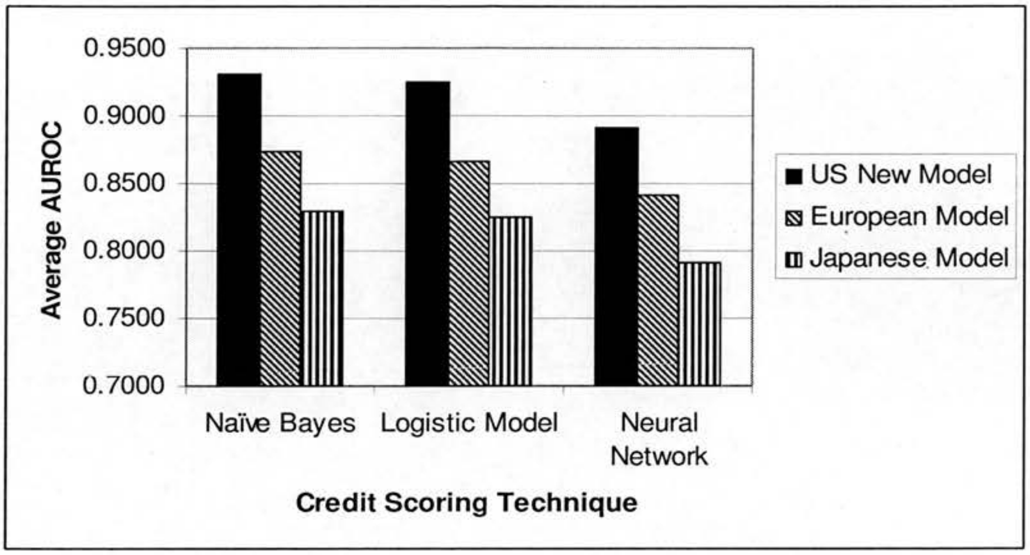
In addition, from Figure 10.8, it is clear that the European and Japanese SMO models do not display any discriminating utility, since the classification power remains constant in most time periods. In contrast, the US SMO model presents a more reliable result.

10.6.2 AUROC Analysis

10.6.2.1 Original Data Comparative Analysis

The following bar chart is employed to conduct a comparative analysis based on average AUROC values:

Figure 10.10 Original Data Comparative Analysis (Average AUROC Value)



Clearly, apart from the credit scoring techniques, the new USA model shows the highest average AUROC value (original dataset), followed by the European model and Japanese model. Furthermore, the difference of the average AUROC value between the USA new model and the Japanese model among three modelling techniques is around 0.1. The result indicates a moderate difference between the best market model and the worst market model in terms of the average AUROC data.

10.6.2.2 Exploring Time Scale

The following line charts can be used to explore the time scale effects among different market models based on the AUROC values.

Figure 10.11 Naïve Bayes Comparative Analysis (AUROC)

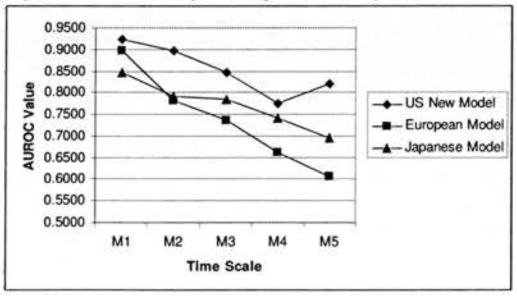


Figure 10.12 Logistic Regression Comparative Analysis (AUROC)

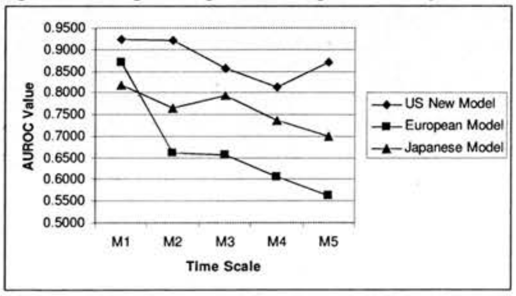
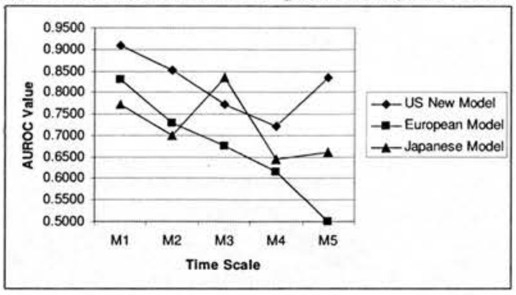


Figure 10.13 Neural Network Comparative Analysis (AUROC)



From Figure 10.11 to Figure 10.13, it is obvious that USA new model has the best performance among most time scales. The only special case is Neural Network model in M3. Another interesting finding is that regardless of the modelling techniques, the European model has better performance than the Japanese model in the year before financial distress, although not in other time scales. Furthermore, the difference among three modelling techniques tends to expand five years before financial distress. For example, the difference between the USA and Japanese Neural Network models is 0.1724 and the difference between the USA and European Neural Network models is 0.3338. As a result, it can be concluded that the performance of the USA new model is still sound five years before financial distress, but the same cannot be concluded for the European and the Japanese market models.

Thus far, from the results in cross-border comparative analysis, it is clear that the new US model displays the best prediction power among all three market models. A possible explanation is that the five key variables are originally selected based on the USA market and hence, the new USA model shows better performance than other market models. However, although the new USA model has better performance, the performance among different market models is very similar in M1. Therefore, it can be concluded that the five key variables still present sound short term prediction power in the European and Japanese markets, although the longer the time period before financial distress, the greater the difference across different markets, in particular in terms of the AUROC value. For example, the US has significantly better AUROC value than Japan or Europe five years prior to financial distress (2000).

10.7 Concluding Remarks

This chapter applied the five key variables: *Debt Ratio*, *Total Debt / (Total Debt + Market Capitalization)*, *Total Assets*, *Operating Cash Flow* and *Government Debt / GD*, which were shown to have sound classification properties in Chapter Eight, to a new US dataset as well as to European and Japanese datasets. An international comparative analysis examining the applicability of these five variables to different markets was carried out.

Using the original dataset, the average accuracy rate was found to be above 86.77% and the average AUROC value above 0.7906 for the new USA model, the European model and the Japanese model. This implies that the five variables have good prediction power among all target markets. When comparing the average performance, the three market models had similar performance in terms of the average accuracy rate, although the USA model comes first.

In time scale analysis, all market models presented good prediction performance in the year before financial distress. However, after M1, the European and Japanese models tended to have weaker prediction power (particularly for AUROC) the longer the time period prior to financial distress. USA new model displayed the best performance in all time scales. Therefore, it can be concluded that the five key variables have good short-term prediction power in all three target markets, but less so for European and Japanese markets on a long term basis.

With regards to the types of error analysis, SMO model showed the best performance to deal with Type II error and the worst ability to manage Type I error among three target markets. This conclusion is consistent with the results in Chapter Eight. However, for the European and Japanese markets, SMO model only has limited discriminating power based on Type I error in most of the time periods. Regarding the ability to deal with Type I error, Naïve Bayes model shows the best performance not only based on the average accuracy rate but also based on each time scale.

Based on the findings in this chapter, the new USA model obtained better results than European and Japanese models. A possible explanation is that the five key variables are originally selected based on the USA market. Furthermore, this also implies that a financial distress model has potentially better prediction ability when based on a single market. Notwithstanding the above, it was decided to explore the performance of a generic global model given that single-market-model construction would be far too time-consuming and costly. This issue of developing a generic model will be discussed in the next chapter.

Chapter ELEVEN

Generic Global Model Development and Performance Evaluation

11.1 Introduction

In Chapter Ten, the USA new model displayed better performance than European and Japanese models in most time scales. It was concluded that a financial distress prediction model has potentially better prediction ability when based on a single market. However, model construction is costly and time consuming. This chapter explores the performance of a generic global model. A composite model is constructed by combining data from US, European and Japanese markets. Overall, a sample of 491 healthy and 68 distressed retail firms is studied over the five-year time period 2000 to 2004 (see Table 10.2).

The next section (Section 11.2) is an evaluation of the prediction ability of the composite model. Section 11.3 then assesses the composite model's practical applicability again through a comparison with Moody's credit ratings. A comparative analysis between the original USA model and the composite model will be presented in Section 11.4. The final section then summarizes the findings in this chapter.

11.2 Prediction Ability Evaluation: Composite Model

As in previous chapters, the *Classification Accuracy Rate* and the *Area under the Receiver Operating Characteristics Curve (AUROC)* were employed to assess the prediction ability of the composite model.

11.2.1 Accuracy Rate Analysis

11.2.1.1 Exploring Time Scale

The results of the accuracy rate analysis based on a five-year time scale are presented in Table 11.1:

Table 11.1 Exploring Time Scale: Accuracy Rate (Composite Model)

Methodology	Performance	M1	M2	M3	M4	M5	Average
Naïve Bayes	Type I Error	44.12%	57.35%	80.88%	60.29%	69.12%	***62.35%***
	Type II Error	4.89%	4.68%	2.44%	8.55%	7.33%	***5.58%***
	Overall	90.34%	88.91%	88.01%	85.15%	85.15%	***87.51%***
Logistic Regression	Type I Error	57.35%	75.00%	83.82%	83.82%	97.06%	***79.41%***
	Type II Error	2.04%	2.65%	1.43%	1.02%	0.41%	***1.51%***
	Overall	91.23%	88.55%	88.55%	88.91%	87.84%	***89.02%***
Neural Network	Type I Error	64.71%	72.06%	82.35%	67.65%	89.71%	***75.30%***
	Type II Error	3.46%	3.67%	3.46%	2.24%	3.05%	***3.18%***
	Overall	89.09%	88.01%	86.94%	89.80%	86.40%	***88.05%***
SMO	Type I Error	97.06%	97.06%	98.53%	100.00%	100.00%	***98.53%***
	Type II Error	0.00%	0.00%	0.00%	0.00%	0.00%	***0.00%***
	Overall	88.19%	88.19%	88.01%	87.84%	87.84%	***88.01%***
Recursive Partitioning	Type I Error	61.76%	85.29%	88.24%	94.12%	100.00%	***85.88%***
	Type II Error	2.85%	3.46%	1.02%	0.20%	0.41%	***1.59%***
	Overall	89.98%	86.58%	88.37%	88.37%	87.48%	***88.16%***

Table 11.1 shows that almost all credit scoring models display the best performance in M1, or, the year before financial distress. The only exception is Neural Network model, which presents the best performance in M4. However, the accuracy rate is very similar across different time periods. For example, taking all modelling techniques together, the biggest performance gap between the M1 and M5 is only around 5.2% (it occurs based on the Naïve Bayes model). This result can be also detected by comparing the average accuracy rate; the largest difference among different credit scoring techniques in this case is only around 1.5%. It can be concluded that time scale effects based on accuracy rate analysis are not obvious. Furthermore, it is hard to say which credit scoring model has the absolute best

performance, since their performances fluctuate according to the different time periods. The performances of the different modelling techniques are also very close.

11.2.1.2 Types of Error

Results similar to those in the previous chapters were found from the types of error analysis. Among the three types of models, the SMO model copes best with Type II error but worst with Type I error. Nevertheless, the result also implies that using the SMO technique, the five variables will not have good classification utility, since SMO model only displays limited discriminating ability to manage the Type I error. Regarding the ability to deal with the Type I error, the Naïve Bayes model continues to show the best performance not only based on the average accuracy rate but also within each time scale.

11.2.2 AUROC Value Analysis

11.2.2.1 Exploring Time Scale

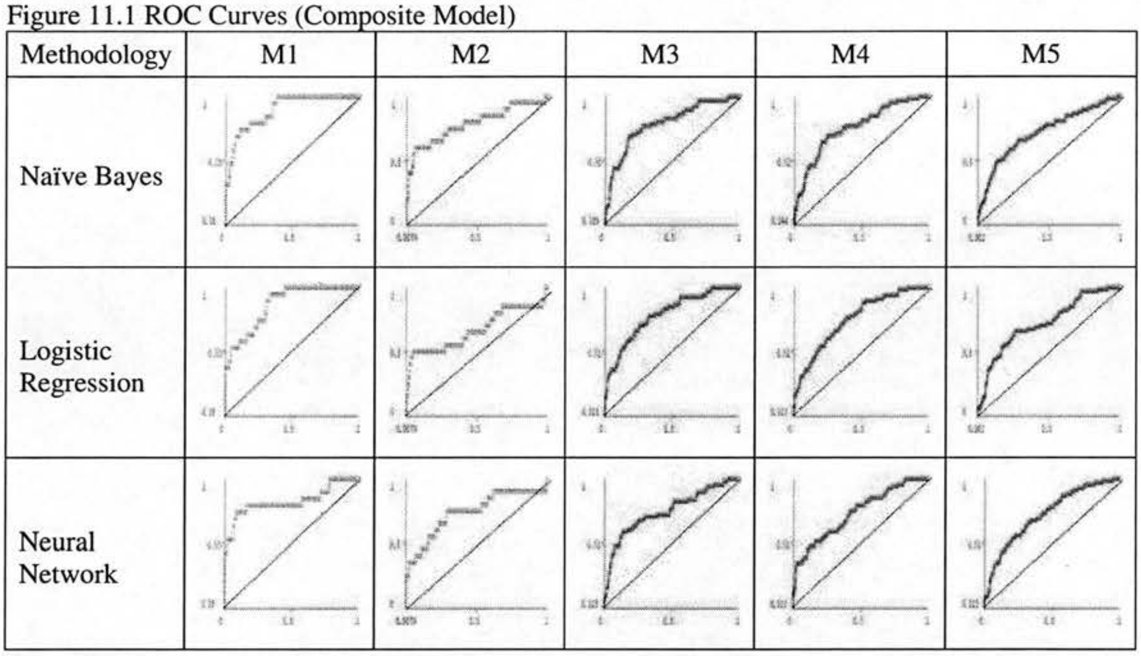
The results of the AUROC analysis based on a five-year time scale are presented in Table 11.2:

Table 11.2 Exploring Time Scale: AUROC Value (Composite Model)

Methodology	M1	M2	M3	M4	M5	Average
Naïve Bayes	0.8781	0.8400	0.7972	0.7649	0.7202	***0.8001***
Logistic Regression	0.8769	0.8300	0.7862	0.7538	0.7203	***0.7934***
Neural Network	0.8472	0.8017	0.7451	0.7363	0.7228	***0.7706***

Clearly, all credit-scoring models show the best performance in M1. Moreover, the AUROC value is above 0.8472 for all of the modelling techniques in M1. The modelling techniques still remain good five years before financial distress, as the AUROC value remains above 0.7202. When comparing the performance among different modelling techniques, Naïve Bayes model shows slightly better performance on average and in most time periods than the other credit scoring models. The only special case is the Neural Network model. This shows the highest AUROC value in M5. The conclusion can be also confirmed by the ROC curves, since the area between the ROC curve and the diagonal line of Naïve Bayes model is slightly larger than the other two modelling techniques from M1 to M4 (see Figure 11.1). Again, the difference in AUROC values among the three credit scoring techniques is very small. For example, the maximum difference of the average AUROC value is below 0.03.

Figure 11.1 ROC Curves (Composite Model)



11.2.3 Concluding Remarks for the Evaluation of Model Prediction Ability

It can be concluded that almost all five credit-scoring techniques have the best classification ability in the year prior to financial distress, with accuracy rates of 88.19% and above as well as AUROC values of 0.8472 and above. (The only exception is Neural Network model in M4 based on the accuracy rate) Furthermore, these techniques still remain good five years before financial distress, with accuracy rates above 85.15% and AUROC values above 0.7202.

No modelling methodology has the absolute best classification ability based on the accuracy rate. The latter tends to vary depending on the time scale considered. For example, Logistic Regression models show best performance in 2004, but not in other time periods. However, Naïve Bayes model presents slightly better performance from the AUROC analysis than other credit scoring models. Finally, if the focus is on the average performance of each modelling technique, then the performance among five credit scoring approaches is very similar. (The maximum difference of the average accuracy rate is only 1.5% and the maximum difference of the AUROC value is only 0.03.)

Thus far, the findings above prove that models have good discriminating ability, even if the time period is five years before financial distress. However, as in Chapter Six, the current results are potentially overly optimistic. A holdout sample has not been used in this research, due to the sample size limits. In order to overcome this problem, the evaluation of the practical applicability through a comparison with Moody's credit rating is conducted and will be discussed in the next section.

11.3 Practical Applicability Evaluation: Composite Model

11.3.1 Moody's Rating

The Logistic Regression, Neural Network and SMO models in the year prior to financial distress were selected for comparison with Moody's credit rating results. As mentioned in Chapter Nine, there are only eight rating grades given (Aa to C) for the retailing industry in Moody's system. Hence, in this study, rating data is ranked according to credit score and divided into eight groups with the same sample size. Moreover, Moody's ratings were only available for a limited number of companies. This is because firms experience the credit rating process only in special circumstances, such as issuing corporate bond abroad. Finally, the sample size for comparative analysis with Moody's ratings is 84 (include 66 USA companies, nine European companies and nine Japanese companies) in each time period. Again, Kolmogorov-Smirnov (K-S) test, Distance analysis, Weighted Kappa analysis and Graphical Bubble charts are employed for comparative analysis.

11.3.2 Kolmogorov-Smirnov (K-S) Test

The Kolmogorov-Smirnov test assesses whether two datasets differ significantly from each other. A p-value greater than 0.05 would indicate that the two samples come from a similar distribution. Results of the K-S Test are shown in Table 11.3.

Table 11.3 Two-sample Kolmogorov-Smirnov (K-S) test (Composite Model)

Methodology	K-S	2004	2003	2002	2001	2000
Logistic Regression	Z Value	1.620	1.852	3.163	4.938	5.401
	p-value	0.010	0.002	0	0	0
Neural Network	Z Value	3.858	3.626	3.009	2.315	1.620
	p-value	0	0	0	0	0.010
SMO	Z Value	1.080	1.620	2.237	2.237	2.469
	p-value	0.194	0.010	0	0	0

The result shows that only SMO has similar rankings to Moody's in 2004.

11.3.3 Distance Analysis

The results of the distance analysis are illustrated in Table 11.4.

Table 11.4 Overall Distances Results (Composite Model)

Methodology	2004	2003	2002	2001	2000	*Average Distance*
Logistic Regression	1.5595	1.7381	1.8095	2.5714	2.9167	***2.11904***
Neural Network	1.8810	1.9286	1.4286	1.2976	1.1786	***1.54288***
SMO	1.3929	1.3929	1.4762	1.5357	1.7381	***1.50716***

From Table 11.4, Logistic Regression model shows the weakest similarity with Moody's, as its average distance is highest amongst the three models. In contrast, SMO model displays the best performance. However, although Neural Network model's average distance is higher than SMO model, the difference is very small. Therefore, it can be concluded that the similarity of SMO model is slightly better than Neural Network model and relatively better than Logistic Regression model.

11.3.4 Weighted Kappa Analysis

The values of weighted Kappa are presented in Table 11.5.

Table 11.5 Weighted Kappa Analysis (Composite Model)

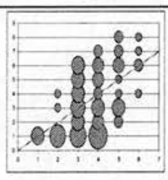
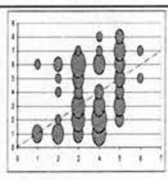
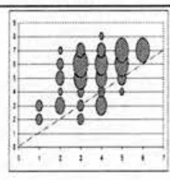
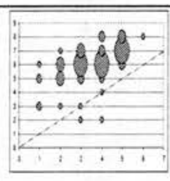
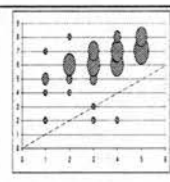
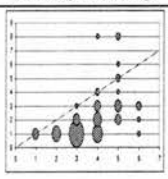
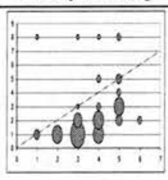
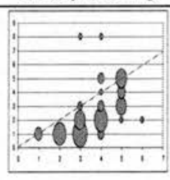
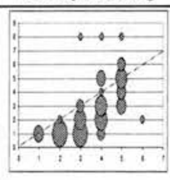
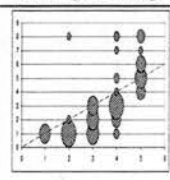
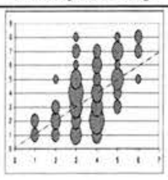
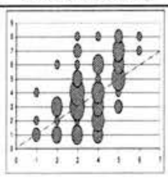
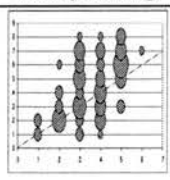
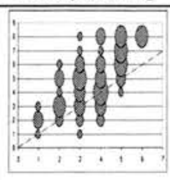
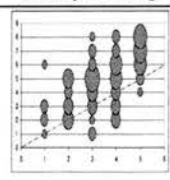
Methodology	2004	2003	2002	2001	2000	*Average Weighted Kappa*
Logistic Regression	0.2381	0.1463	0.1208	0.0339	0.0223	***0.11228***
Neural Network	0.1367	0.1069	0.2512	0.3280	0.4208	***0.24872***
SMO	0.2814	0.2819	0.2774	0.2473	0.1998	***0.25756***

Unsurprisingly, the results are similar to those from distance analysis. Average weighted Kappa results suggest that SMO is the better performing model amongst the three models, closely followed by Neural Network. Logistic Regression still shows lowest performance in terms of agreement with Moody's.

11.3.5 Graphical Bubble Charts Analysis

The bubble charts of the composite model are presented in Figure 11.2:

Figure 11.2 Bubble Charts of the Composite Model

Methodology	2004	2003	2002	2001	2000
Logistic Regression Model					
	Moody's Rating	Moody's Rating	Moody's Rating	Moody's Rating	Moody's Rating
Neural Network					
	Moody's Rating	Moody's Rating	Moody's Rating	Moody's Rating	Moody's Rating
SMO					
	Moody's Rating	Moody's Rating	Moody's Rating	Moody's Rating	Moody's Rating

Out of the three modelling approaches, Neural Network models show the weakest similarity to Moody's in 2004. Indeed, only a few bubbles are close to the diagonal line and most large size bubbles are away from the diagonal line. Logistic Regression model's bubble chart in 2000 appears worse than the results for Neural Network model in 2004.

Conclusions from these bubble charts confirm results from distance and weighted Kappa analyses in sections 11.3.4 and 11.3.5. In 2004, the distance value from Neural Network approach is 1.881 (highest among three models) and weighted Kappa value is 0.1367 (lowest among three models). The situation is indeed worse for Logistic Regression model in 2000, since the distance value is 2.9167 and weighted Kappa value is 0.0223.

In addition, the similarity of research models to Moody's can also be analysed over time. Based on the bubble charts, the performance of Logistic Regression model improves year by year from 2000 to 2004, as more large size bubbles are increasingly concentrated on the diagonal line. The opposite occurs for Neural Network model in the same time period. Compared with the trends of the other two credit scoring techniques, the SMO model shows a more consistent performance between 2000 and 2004.

Another interesting finding is that for all credit scoring techniques, the bubbles tend to move downwards year by year from 2000 to 2004. As mentioned in Chapter Nine, bubbles above the diagonal line indicate better ratings for Moody's than for research models. Bubbles below the diagonal line indicate lower rating for Moody's than for research models. Thus, adopting Moody's as a benchmark, it can be said that research models possibly underrate the credit situation of sample companies in 2000 and overrate the credit situation in 2004.

11.3.6 **Concluding Remarks for the Evaluation of Practical Applicability**

Based on the Kolmogorov-Smirnov significance test, distance measure, and weighted Kappa measure, SMO models performed best, followed closely by the

Neural Network model. Logistic Regression models showed weakest performance in terms of similarity with Moody's. The bubble chart analysis proved extremely useful not only for understanding the similarity between two ordinal datasets, but also for detecting model performance trends.

It was found that the performance of Logistic Regression model improved the closer it got to the year of financial distress, whilst the opposite occurred for Neural Network model. Comparing with the trends of the other two credit scoring techniques, SMO model showed a more consistent performance between 2000 and 2004. Furthermore, from the Bubble Chart analysis, it can be concluded that the research models possibly underrate the credit situation of sample companies in 2000 and overrate the credit situation in 2004.

Thus far, this research has evaluated the prediction power and the practical applicability of the composite model. The composite model has good prediction power not only in the short term, but also in the long run. Does the composite model have better performance than the original US prediction model? The next section will focus on a comparative analysis between the original USA model and the composite model.

11.4 Comparative Analysis: Original USA Model and Composite Model

11.4.1 Comparative Analysis of Model Prediction Power

11.4.1.1 Accuracy Rate Analysis

The results of the accuracy rates between the original model and the composite model are illustrated in the following line charts:

Figure 11.3 Naïve Bayes Comparative Analysis (Accuracy Rate)

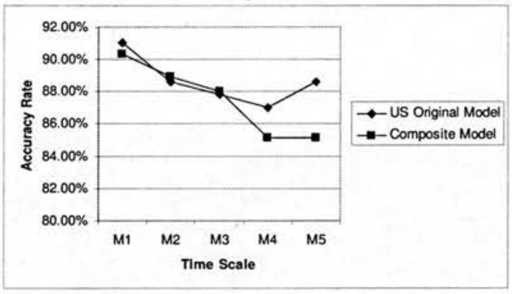


Figure 11.4 Logistic Regression Comparative Analysis (Accuracy Rate)

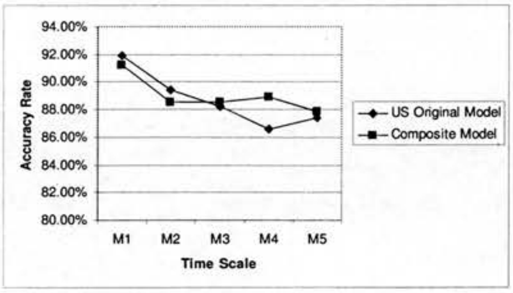


Figure 11.5 Neural Network Comparative Analysis (Accuracy Rate)

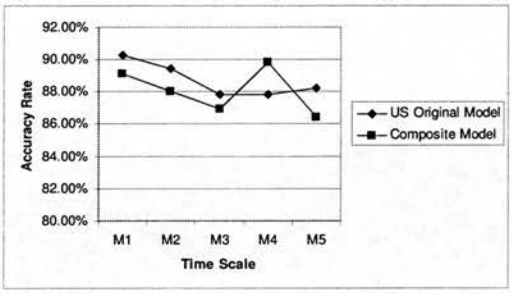


Figure 11.6 SMO Comparative Analysis (Accuracy Rate)

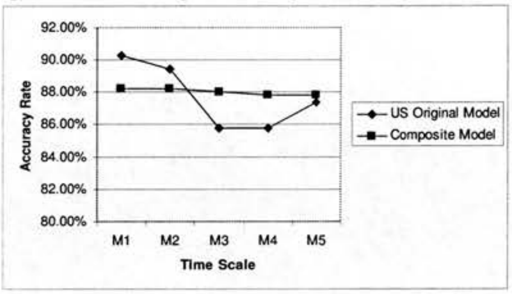
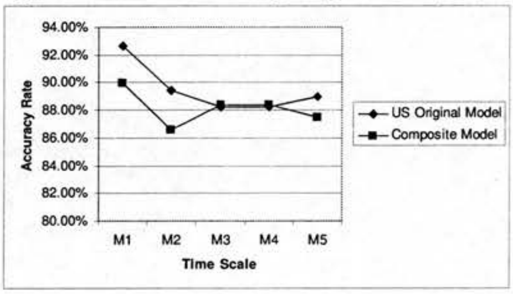


Figure 11.7 Recursive Partitioning Comparative Analysis (Accuracy Rate)



It appears difficult to conclude which model has better performance from the line charts; the comparative performance of each model varies in terms of different time scales and modelling techniques. Performance comparison based on average accuracy rates (see Table 11.6) is also inconclusive.

Table 11.6 Average Accuracy Rate Comparison

	Naïve Bayes	Logistic Regression	Neural Network	SMO	Recursive Partitioning
USA Original Model	88.62%	88.70%	88.70%	87.72%	89.51%
Composite Model	87.51%	89.02%	88.05%	88.01%	88.16%
Difference	1.11%	-0.32%	0.65%	-0.29%	1.35%

It seems that two models have very similar performance. In fact, the largest difference between these two models is a mere 1.35%. Furthermore, in all of the modelling techniques, the average accuracy rate is above 87.51%. This indicates that both models have sound prediction ability. In other words, the five variables are good predictors for forecasting retail default.

11.4.1.2 AUROC Value Analysis

Again, line charts can be employed to compare the AUROC values of the original and composite models:

Figure 11.8 Naïve Bayes Comparative Analysis (AUROC)

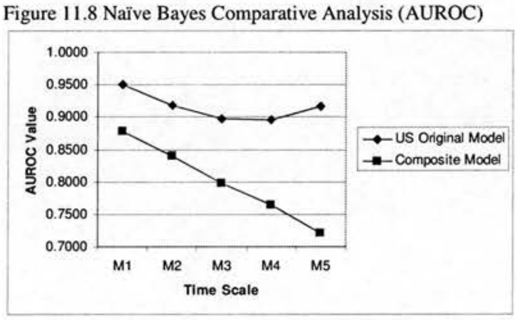


Figure 11.9 Logistic Regression Comparative Analysis (AUROC)

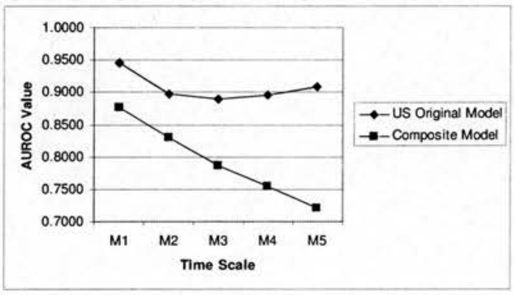
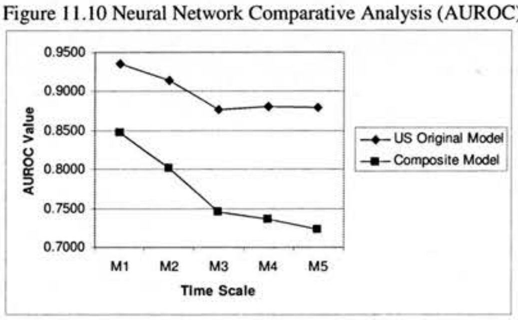


Figure 11.10 Neural Network Comparative Analysis (AUROC)



Figures 11.8 to 11.10 show that the AUROC performance of the USA original model is better than the composite model, regardless of modelling techniques or the time scales. Moreover, the difference becomes larger the longer the period before financial distress.

11.4.2 Practical Applicability Comparative Analysis

11.4.2.1 Distance Analysis

Figure 11.11 Logistic Regression Comparative Analysis (Distance)

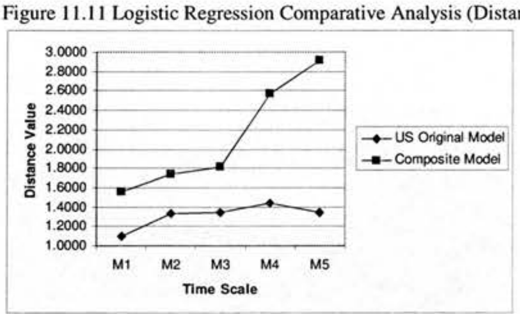


Figure 11.12 Neural Network Comparative Analysis (Distance)

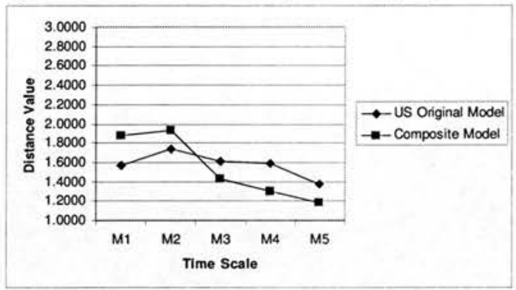
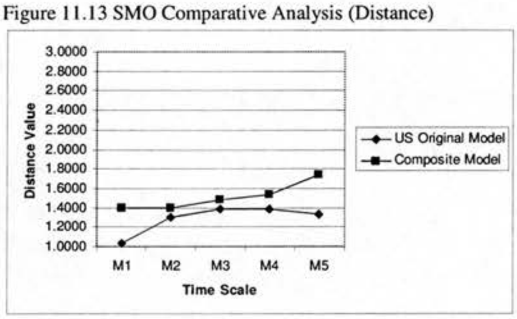


Figure 11.13 SMO Comparative Analysis (Distance)



In terms of distance, the line charts in Figures 11.11, 11.12, and 11.13 show higher values for the composite model than the USA original model for Logistic Regression and SMO techniques. Moreover, the difference between the two models is larger for Logistic Regression than for SMO. The two models show similar performance for the Neural Network technique.

Using distance analysis, it can be concluded that the USA original model has relatively greater similarity with Moody's rating than the composite model when using the Logistic Regression, and slightly better similarity with Moody's rating than the composite model when using the SMO technique. With regards to the Neural Network technique, the two models display similar performance to Moody's rating.

11.4.2.2 Weighted Kappa Analysis

The following line charts are used to conduct the comparative analysis based on weighted Kappa:

Figure 11.14 Logistic Regression Comparative Analysis (Weighted Kappa)

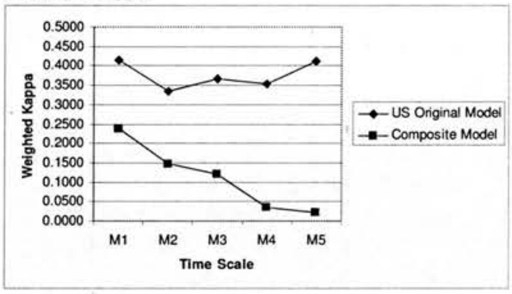


Figure 11.15 Neural Network Comparative Analysis (Weighted Kappa)

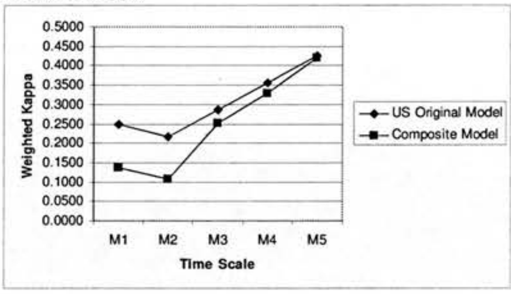
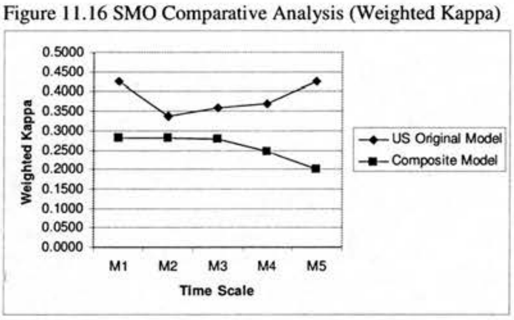


Figure 11.16 SMO Comparative Analysis (Weighted Kappa)



The line charts show that in all of the modelling techniques, the US original model has higher weighted Kappa, and therefore greater agreement with Moody's, than the composite model. Moreover, for the Logistic Regression and SMO techniques, the difference between the USA original model and the composite model becomes larger the longer the time period before financial distress. This was also found above in distance analysis comparisons. In contrast, for the Neural Network technique, although the USA original model shows better performance than the composite model, the difference becomes smaller the longer the time period prior to default. This latter finding is in contrast with the distance analysis comparisons above.

11.5 Concluding Remarks

In this chapter, a composite model was constructed by combining data from US, European and Japanese markets with the goal of assessing the performance of a generic global model. Overall, a sample of 491 healthy and 68 distressed retail firms was studied over a five-year time period from 2000 to 2004.

Results on the prediction ability of the composite model indicated that all five credit-scoring techniques have the best classification ability in the year prior to the financial distress, with accuracy rates of 88.19% and above, and AUROC values of 0.8472 and above. (The only exception is Neural Network model in M4 based on the accuracy rate) Furthermore, these techniques still remained sound five years before financial distress; the accuracy rate is above 85.15% and AUROC value is above 0.7202.

With variations in accuracy rate performance for each period, it was hard to conclude which modelling methodology had the absolute best classification ability. Based on AUROC analysis, Naïve Bayes model presented slightly better performance than other two credit scoring models, but the difference among modelling techniques was not significant.

Finally, regarding the analysis of types of error, results similar to those in Chapter Ten were reached: the five variables did not present good classification utility based on the SMO technique. Indeed, the SMO models displayed limited discriminating ability to manage the Type I error. Naïve Bayes model still showed the best performance to deal with Type I error not only based on the average accuracy rate but also within each time period.

On the topic of each model's practical applicability, it can be concluded that SMO model is the better performing model amongst the three modelling techniques, although it is closely followed by Neural Network model. Logistic Regression model appeared to be least similar to Moody's. From bubble chart analyses, it was found that the performance of Logistic Regression model improved year by year from 2000 to 2004, whilst the opposite occurs for Neural Network model in the same time period. Moreover, the research models possibly underrated the credit situation of sample companies in 2000 and overrated the credit situation in 2004.

So far, the conclusions indicate a paradoxical result (as in Chapter Nine). Logistic Regression model and Neural Network model display slightly better classification ability than SMO model in the year before default, but SMO model seems to be stronger in terms of comparability with Moody's rankings. A possible explanation is that the Logistic Regression and Neural Network model fit the sample too closely, hence overfitting occurred. For SMO, this did not occur.

In connection with the comparative analysis between the original US model and the US-Europe-Japan composite model, it was found that the composite model has similar performance to the US original model in terms of the accuracy rate analysis. However, the same cannot be concluded for AUROC analysis. The results also showed that the US original model has higher AUROC values, regardless of the modelling technique or time scale in question.

Finally, the US original model appeared to have better practical applicability than the composite model when compared with Moody's ratings. This was the case for both distance and weighted Kappa measures and for almost all modelling techniques. The only special case was the Neural Network technique based on distance measures. Here, the composite and original US models displayed similar performance.

In summary, it seems that the USA original model has better performance than the composite model not only based on prediction ability, but also based on practical applicability. This confirms that a default prediction model has potentially better prediction ability when based on a single market. However, model construction is time-consuming and costly. Hence, global model development was still an important direction for future research. In this research, the composite model was only based on US, European and Japanese markets. More world retail markets can be included for future studies in order to ensure the model's prediction utility and practical applicability.

Part Six

Conclusions and Discussions

- Chapter Twelve: ***Conclusions and Discussions***

  Chapter Twelve summarizes the findings, outlines the limitations and suggests possible future directions for research in the corporate performance measurement and default prediction domain. In addition, it illustrates the contributions to interested parties in this research area.

Chapter TWELVE

Conclusions and Discussions

12.1 Summary of Research Findings

The measurement and forecasting of corporate performance is of critical importance not only to mangers but also to external stakeholders of the company such as investors, lenders and policy makers. These players seek better methods of performance measurement and prediction to aid optimal decision-making. From a review of previous studies, it was found that a gap exists between the studies in performance measurement systems and performance prediction models.

On the one hand, performance measurement systems are founded on a theoretical framework that concentrates on a general business context rather than a specific industrial sector (Ansell, 1992). Few empirical studies examine the prediction ability of current performance measurement systems despite the importance of performance prediction for decision makers (Smith, 2005). On the other hand, default prediction models have the ability to provide a practical platform for forecasting company performance. However, most previous default prediction models lack theoretical framework and realistic view for variable selection.

Drawing on above, the primary objective of the research has been to ***develop effective performance measurement and prediction models*** using a wide range of potential performance variables and credit-scoring techniques, in order to fill the gap between the previous studies. As mentioned in Chapter One, this primary objective was divided into three sub-objectives:

- Developing a Corporate Performance Measurement Framework
- Developing Financial Distress Prediction Models
- Evaluating Model Utility

In this research, the retailing industry was selected to illustrate the workings of the model for retail risk assessment and evaluation increasingly be a critical area of research (Dawson, 2000). More specifically, USA's retailing sector was chosen because of the clear definitions and reporting of financial distress. The following sections briefly outline how the three sub-objectives have been achieved.

12.1.1 Developing a Corporate Performance Measurement Framework

Given the gaps in the literature, the corporate performance framework needed to incorporate professional views from industry and possess theoretical grounding. Regarding the latter point, Hunt's (2000) *Resource-Advantage (R-A) Theory of Competition* was chosen based on its strengths. For example:

- The fundamental premises of R-A theory are closer to the reality and are highly related to the retail environment.
- R-A theory can explain the dynamic, evolutionary and disequilibrium process of competition in retail.
- R-A theory provides a more complete blueprint for research framework construction, as it considers corporate competition advantage based on both internal resources and external influences.

With regards to the need to incorporate professional views from industry, the listing of potential performance measures were not limited to the existing literature, but also included practitioners' viewpoints. The researcher carried out 25 interviews covering three stakeholder groups: the retail companies' management team, the banks business loan department team, and the investment institutions' industrial analysts' team. Interestingly, members of these three groups shared a common viewpoint in most cases. The most obvious difference among stakeholders was their roles vis-à-vis retail performance evaluation. Overall, 170 retail performance measures were found and divided using R-A theory into: *Internal Resource Group* and *External Factors Group*.

The interviews complemented previous studies on the potential performance measures. However, they did not clarify the importance of the performance measures. Consequently, the researcher also carried out a survey in order to obtain more insights from the research context. The researcher envisaged the possibility of eliminating current variables if they proved unimportant after conducting the survey. Therefore, the survey was viewed as a robust examination of the model framework.

44 variables were selected for final e-questionnaire design and 435 e-questionnaires were sent out to US, UK and Taiwan. Overall, 151 responses were received giving an overall response rate of 34.71%. The descriptive analysis showed that the 44 variables could be considered in the research framework. From the comparative analyses among different retail formats and countries, the results did not indicate any difference in terms of variable importance. However, the importance of the 44 variables was found to be statistically different between departments. The primary reason for dissimilarity was that departments tended to have different priorities or viewpoints regarding performance.

Finally, a case study was carried out in order to investigate the difference between the 'expected value' and the 'actual performance' of selected variables. Overall, the viewpoints on 'expected value' and 'actual performance' were statistically different. This could be explained by the target company's policy to set high annual objectives in order to encourage employees to outperform themselves. The gap between expected and actual performance tended to be significant as a matter of course.

12.1.2 Developing Financial Distress Prediction Models

Credit-scoring techniques classify whether a business is likely to have high or low risk of, say, default based on information gleaned from business including its history. In the current study, credit-scoring techniques were used to differentiate 'healthy' from 'distressed' firms according to a specific definition. For example a 'healthy' firm may be a company that will not default or become bankrupt in the next year. Credit-scoring analyses help corporate performance predictions by producing scores that are related to the companies' probability of default or bankruptcy.

Five credit scoring techniques: Naïve Bayes, Logistic Regression, Recursive Partitioning, Artificial Neural Network, and Sequential Minimal Optimization (SMO), were employed on a dataset of 195 healthy and 51 distressed US retailers with 67 variables over five time periods: 1994-1998, 1995-1999, 1996-2000, 1997-2001 and 1998-2002. Ideally, a limited number of key variables or principal components should be selected in order to produce a relatively simple model for predicting default. The quality of the final key variables was ensured through time-scale consideration, outlier elimination and preliminary univariate analysis.

This research also carried out a cross-validation process against potential overfitting problems. The 10-folders cross-validation method was chosen, as this approach only wasted 10% of total data and the training cost was much lower than the leave-one-out method. Given that this research considered two variable selection methods (*Forward Stepwise Approach* and *Principal Component Analysis*), two different variable groups, five different time periods, and five credit scoring modelling techniques, a total number of 100 models were constructed.

12.1.3 Model Utility Evaluation

The default prediction model was evaluated in terms of both model prediction power and practical applicability. *Classification Accuracy Rate* and *the Area under the Receiver Operating Characteristics curve (AUROC)* were employed to evaluate the model prediction power. Practical applicability of the model was assessed by two approaches. First, rankings from the study were compared with those from a standard rating system—in this case the well-established *Moody's Credit Rating*. It was assumed that the higher the degree of similarity between the two sets of rankings, the greater the credibility of the default prediction model. Another method was to apply the original model to different markets.

There was sufficient evidence that the five credit scoring techniques have sound classification ability in the time period of one year before financial distress regardless of the variable selection approach used. Moreover, they remained sound even five years prior to financial distress; the classification accuracy rates exceeded 78.46% and AUROC values exceeded 0.7345. These numbers indicated that the developed models were theoretically sound. However, each modelling technique performed differently over time depending on the variable selection approach. It was therefore difficult to pinpoint which technique had the best classification ability. This is consistent with most previous studies addressed in Chapter Two. In addition, the analysis showed weak external environment influences on default assessment for all five credit-scoring techniques. Finally, almost all Forward Stepwise models possessed accuracy rates and AUROC values that were higher than PCA models'. This was most obvious one, two and three years before default.

A comparative analysis with Moody's to assess the practical applicability of the models was carried out using four techniques: Kolmogorov-Smirnov (K-S) test, Distance analysis, and Weighted Kappa analysis and Graphical Bubble charts. Logistic Regression and SMO models were found to be most comparable with Moody's. Furthermore, as found during model prediction ability assessment, Forward Stepwise models still outperformed PCA models based on both distance and the weighted Kappa analyses. Having a similar conclusion to that obtained during model prediction ability assessment indicate that the original research idea to compare with Moody's rating was reasonable. It also proves that the five variables: *Debt Ratio*, *Total Debt / (Total Debt + Market Capitalization)*, *Total Assets*, *Operating Cash Flow* and *Government Debt / GDP*, which were selected by the Forward Stepwise approach, have sound prediction performance in the USA market.

However, the findings were paradoxical in that the Logistic Regression and Neural Network stepwise models both produced slightly better prediction results than the SMO stepwise models, whilst SMO models outperformed the Neural Network model (to a large extent) as well as the Logistic Regression models (to a small extent) in terms of company ranking. This can be explained by possible overfitting in both Neural Network and Logistic Regression stepwise models but not in the SMO models.

On the topic of international comparison, the five key variables were applied to a new US data set as well as the European and Japanese markets. The model ensuing from the new US dataset presented better performance than European or Japanese models. This is probably because the five key variables were originally selected based on the USA market. The results also imply that a financial distress prediction model has potentially better prediction ability when based on a single market.

With regards to the types of error analysis, SMO model showed the best performance to deal with the Type II error and the worst ability to manage the Type I error in all area models. However, for the European and Japanese markets, SMO model does not have any discriminating power based on the Type I error in most of the time periods studied. It was concluded that the five variables only have limited classification utility when using the SMO technique in the European and Japanese markets.

A high Type I error also indicates that most sample companies are classified as healthy companies and it will damage the benefits from some interested parties. For example, Type I error may cause an investor to lose the entire investment, while Type II error may only cause an investor to lose the potential dividends or capital gains. With regards to the ability to deal with the Type I error, Naïve Bayes model showed the best performance for all markets, not only in terms of average accuracy rate but also in each time scale.

Following the above, it was decided to explore the possibility of creating a generic global model, as, in practice, continuous single market model construction would tend to be time-consuming and costly. A composite model was constructed by combining data from US, European and Japanese markets. This composite model generated sound predictions, even up to five years before financial distress. The accuracy rates exceeded 85.15% and AUROC values exceeded 0.7202. Moreover, the composite model has similar prediction performance as the original US model in terms of accuracy rates. However, it presented a worse prediction utility than the original US model in terms of AUROC values.

On the topic of composite model's practical applicability, it can be concluded that SMO model is the better performing model amongst the three, although it is closely followed by Neural Network model. Logistic Regression model appeared to be least similar to Moody's. By comparing the performance with the US original model, it was found that the US original model appeared to have better practical applicability than the composite model when compared with Moody's ratings. This was the case for both distance and weighted Kappa measures and for almost all modelling techniques. It seems that the USA original model has better performance than the composite model not only based on prediction ability, but also based on practical applicability. This confirms that a default prediction model has potentially better prediction ability when based on a single market.

Drawing on the discussions above, it can be concluded that the default prediction model in this research has good performance in terms of both prediction utility and practical applicability for single markets and the composite market. This confirms usefulness of having an underlying research framework and fundamental theory. Furthermore, it indicates that the theoretical groundwork and realistic view for variable selection are critical in financial distress prediction. Overall, the primary research objective, ***Developing an Effective Corporate Performance Measurement and Prediction System*** was achieved satisfactorily in this research.

12.2 Research Contributions

A sound performance measurement and prediction system is of interest to several audiences, including academics, business community and policy-makers. The impacts of this research on these interested parties are outlined below.

12.2.1 Contributions for Academics

This research considers several key issues, which are different from most previous default prediction studies. The lack of theoretical groundwork for variable selection is a common situation in most default prediction studies. The present research incorporated a much larger number of variables and developed a theoretical framework based on Hunt's (2000) Resource-Advantage (R-A) Theory of Competition. The results indicated that the R-A theory could provide a sound theoretical foundation to assess retail performance. The final default prediction models displayed sound ability in classifying retailing companies into healthy and distressed groups. Moreover, R-A theory provides a platform to fill the gap in the previous studies between both performance measurement systems and default prediction models by overcoming their relevant limitations.

Financial distress researchers often do not go beyond variables mentioned in previous studies. Obviously, such a variable selection method is limited. One of the contributions in this research is the use of a wider range of variables. For example, most previous studies only employ financial ratios to develop financial distress models, such as Beaver (1966), Altman (1968) and Ohlson (1980). In contrast, this research considers more aspects in relation to variable selection to construct default prediction model, such as financial ratios, industrial factors, external environmental factors and cash flow structure. For example, most previous studies only select macroeconomic variables to reflect the influences from external environment. In this research, political, social-cultural and technological variables were also selected. Drawing on the above, it can be argued that most variables in previous studies have been covered by this research based on the R-A theory framework.

A comparison of five methods for predicting company default: Naïve Bayes, Logistic Regression, Recursive Partitioning, Artificial Neural Network, and Sequential Minimal Optimization (or SMO, a form of Support Vector Machine) was made, bringing insight to which was the most effective and why. The application of the SMO is an innovation in forecasting business financial distress and predicting business default (although Support Vector Machine has been popular in the consumer credit scoring domain). A comparison among different modelling techniques has also been made between local and global models based on comparisons among American, European and East Asian data.

Another aspect of this research also was the incorporation of subjective views from industry players. Through interviewing and surveying a number of stakeholders, qualitative measures of performance were found. Given the importance of developing techniques that can be easily interpreted, the researcher also introduced graphical representations of performance prediction using bubble charts.

12.2.2 Contributions for the Business Community

The constructed financial distress prediction model can be viewed as a decision-making benchmark for stakeholders in the business community. Retail management may use the model to evaluate and so modify a firm's strategy to avoid financial distress. It is also a key point of reference for credit rating, especially when a retail company plans to issue corporate bonds. Lenders (e.g. bank loan departments) can use this model to evaluate a firm's default probability in order to make the appropriate loan decision.

Investors, such as in hedge funds, can obtain default probability from the model and use it as a benchmark to develop an investment strategy. Auditors can use this model to assess a company's vulnerability and avoid the lawsuits arising from the failure to reveal the probability of financial distress. Employees or labour union can obtain information from this model to understand the potential threats to continued employment. Suppliers can use this model to measure the need to terminate a contract with the retail company.

Overall, this research has strong end-user application. It is of interest to the whole industry, not only to selected companies. The whole retail community can therefore benefit from this research. Such benefits can also be extended to other industries.

12.2.3 Contributions for Policy-makers

This research focuses on financial distress prediction in retailing and hence, provides valuable information for governmental or other institutional policy-makers to establish policy on retailing. For example, policy-makers can use the model to obtain information about different retail sectors for budget allocation purposes. Moreover, the research model can also assist government to establish certain regulations such as the publicly listed requirements for retailing.

12.3 Future Research Plans

In parallel to this work, the researcher has explored the use of qualitative variables in the assessment of business performance. This has been achieved through the development of a survey instrument, which was administered to a range of expert stakeholders in UK, USA and East Asia. The incorporation of survey data into the default prediction model was contemplated using Bayesian techniques. Work on qualitative variables as a means of predicting financial distress is still at a preliminary stage. Obviously the prime concern has been obtaining qualitative information from distressed companies. There was greater likelihood to explore companies that are *'close'* to being distressed rather than truly distressed. Thus, building a Bayesian graphical model with both qualitative and quantitative variables to predict financial distress in retailing will be a potential research direction.

In addition, due to the limitations regarding the length of survey, this research only considered 44 variables to be the main performance measures in the e-questionnaire. The survey may be extended by considering more variables in future studies. The research findings also showed that Forward Stepwise models had better classification ability than PCA models. However, the chosen principal components may not have sufficient explanatory power vis-à-vis total variance. Therefore, future research on explained variance can be carried out to enhance the performance of PCA model.

A generic global model development is another important direction for future research. In this research, the original US model was found to outperform the composite model in terms of both model prediction ability and practical applicability. Although this means that a default prediction model has potentially better prediction ability when based on a single market, it was noted that single-market model construction tended to be too time-consuming and costly in practice. In this study, a composite model was constructed, but only based on US, European and Japanese market data. More world retail markets could be included in future studies so as to ensure prediction utility and practical applicability of the default prediction models.

As mentioned in Chapter Two, market-based models, such as Moody's KMV, are also an important technique for predicting financial distress in context. Hence, evaluating the model's practical applicability by comparing credit scoring results with Moody's KMV model can be another future research plan. Finally, it must be noted that the scope of this study was limited to publicly listed firms and the retail market. The study can be extended to non-listed firms as well as other industrial sectors in order to investigating the differences among different aspects.

Part Seven

References

REFERENCES

1. Aaker, D.A (1991) *Managing Brand Equity: Capitalising on the Value of a Brand Name*, Free Press, New York

2. Aaker, D.A (1992) The Value of Brand Equity, *Journal of Business Strategy*, 13, 4, 27-32

3. Altman, E.I. (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance*, 23, 4, 589-609.

4. Altman, E.I. (1983) *Corporate Financial Distress: A Complete Guide to Predicting Avoiding and Dealing with Bankruptcy*, John Wiley & Sons Press, New York

5. Altman, E.I. (1995) Prediction financial distress of companies: revisiting the Z-score and ZETA models, Paper presented at the *Korean Institute of Finance Conference* on Bank Credit Risk management, 28-30 June, Seoul, Korea

6. Altman, E.I., Haldeman, R.G. and Narayanan, P. (1977) Zeta analysis: A new model to identify bankruptcy risk of corporations, *Journal of Banking and Finance*, 1, 1, 29-54

7. Altman, E.I., Hartzell, J. and Peck, M. (1995) *Emerging Markets Corporate Bonds: A Scoring System*, Salomon Brothers, New York

8. Altman, E.I., Marco, G. and Varetto, F. (1994) Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience), *Journal of Banking & Finance*, 18, 3, 505-529

9. Ansell, J. (1992) *Reliability: Industry Risk Assessment*, in Ansell, J. and Wharton, F. (1992) *Risk: Analysis , Assessment and Management*, John Wiley & Sons Ltd., Chichester, England

10. Amir, E. and Lev, B. (1996) Value-relevance of nonfinancial information: The wireless communications industry, *Journal of Accounting and Economics*, 22, 1-3, 3-30

11. Aziz, A., Emanuel, D.C. and Lawson, G.H. (1988) Bankruptcy prediction—an investigation of cash flow based models, *Journal of Management Studies*, 25, 5, 419-437

12. Babin, B.J., Griffin, M. and Boles, J.S. (2004) Buyer reactions to ethical beliefs in the retail environment, *Journal of Business Research*, 57, 10, 1155-1163

13. Baldwin, J. and Glezen, G.W. (1992) Bankruptcy prediction using quarterly financial statement data, *Journal of Accounting, Auditing and Finance*, 7, 3, 269-285

14. Banker, R.D., Potter, G. and Srinivasan, D. (2000) An empirical investigation of an incentive plan that includes nonfinancial performance measures, *Accounting Review*, 75, 1, 65-92

15. Beaver, W.H. (1966) Financial ratios as predictors of failure, *Journal of Accounting Research*, Supplement, 4, 3, 71-111

16. Becchetti, L. and Sierra, J. (2003) Bankruptcy risk and productive efficiency in manufacturing firms, *Journal of Banking and Finance*, 27, 11, 2099-2120

17. Begley, J., Ming, J. and Watts, S.G. (1996) Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models, *Review of Accounting Studies*, 1, 4, 267-284

18. Behn, B.K. and Riley Jr., R.A. (1999) Using nonfinancial information to predict financial performance: The case of the U.S. airline industry, *Journal of Accounting, Auditing and Finance*, 14, 1, 29-56

19. Betts, J. and Belhoul, D. (1987) The effectiveness of incorporating stability measures in company failure models, *Journal of Business, Finance and Accounting*, 14, 3, 323-334

20. Blum, M. (1974) Failing company discriminant analysis, *Journal of Accounting Research*, 12, 1, 1-25

21. Boritz, J. and Kennedy, D. (1995) Effectiveness of neural network types for prediction of business failure, *Expert System with Applications*, 9, 4, 504-512

22. Bowen, R.M., Daley, L.A. and Huber Jr., C.C. (1982) Evidence on the existence and determinants of inter-industry differences in leverage, *Financial Management*, 11, 4, 10-20

23. Brown, M., Grundy, W., Chistianini, N., Lin, D., Sugnet, C., Furey, T., Ares Jr., M., and Haussler, D. (2000) Knowledge-based analysis of microarray gene expression data using support vector machines, *Proceedings of the National Academy of Sciences*, 97, 1, 262-267

24. Bromwich, M. (1988). Managerial accounting definition and scope-from a managerial view. *Management Accounting*, 66, 8, 26-27

25. Bromwich, M. and Bhimani, A. (1989) *Management Accounting: Evolution not Revolution*, The Chartered Institute of Management Accountants, Black Beer Press Limited, Cambridge, England

26. Bryman, A. (2001) *Social Research Methods*, Oxford University Press, Oxford, England

27. Burt, S. and Sparks, L. (2003) Competitive analysis of the retail sector in the UK, *Report submitted to the department of trade and industry*, University of Stirling

28. Buttle, F. (1992) Shopping motives constructionist perspective, *Service Industries Journal*, 12, 3, 349-367

29. Caouette, J.B., Altman, E.I. and Narayanan, P. (1998) *Managing Credit Risk: The Next Great Financial Challenge*, John Wiley & Sons Inc., New York

30. Carson, J.M. and Hoyt, R.E. (1995) Life insurer financial distress: classification models and empirical evidence, *Journal of Risk and Insurance*, 62, 4, 764-775

31. Casey Jr, C.J. (1980) Variation in accounting information load: The effect on loan officers' predictions of bankruptcy, *Accounting Review*, 55, 1, 36-49

32. Casey, C.J. and Bartczak, N.J. (1984) Cash flow—it's not the bottom line. (cover story), *Harvard Business Review*, 62, 4, 61-66

33. Casey, C.J. and Bartczak, N.J. (1985) Using operating cash flow data to predict financial distress: some extensions, *Journal of Accounting Research*, 23, 1, 384-401

34. Chandler, Jr. A.D. (1990) *Scale and Scope: The Dynamics of Industrial Capitalism*, Harvard University Press, USA

35. Chang, Y.H. (2002) *Analysis of Competition in the Convenience Store Industry in Taiwan based on Resource-Advantage Theory*, Doctor of Philosophy Dissertation, University of Edinburgh

36. Charitou, A., Neophytou, E. and Charalambous, C. (2004) Predicting corporate failure: empirical evidence for the UK, *European Accounting Review*, 13, 3, 465-497

37. Chen, K.H. and Shimerda, T.A. (1981) An empirical analysis of useful financial ratios, *Financial Management*, 10, 1, 51-60

38. Chenet, P. and Johansen, J.I. (1999) *Beyond Loyalty: The Next Generation of Strategic Customer Relationship Management*, Oak Tree Press, Ireland

39. Christianini, N. and Shawe-Taylor, J. (2000) *An Introduction to Support Vector Machines*, Cambridge University Press, Cambridge, UK

40. Chou, W.S. (2002) *Multivariate Statistical Analysis with Application of SAS/SPSS*, BestWise Co Ltd., Taiwan (in Chinese)

41. Coats, P.K. and Fant, L.F. (1993) Recognizing financial distress patterns using a neural network tool, *Financial Management*, 22, 3, 142-155

42. Collin-Dufresne, P., Goldstein, R.S. and Martin, J.S. (2001) The determinants of credit spread changes, *Journal of Finance*, 56, 6, 2177-2207

43. Comrey, A.L. and Lee, H.B. (1992). *A First Course in Factor Analysis*, Lawrence Erlbaum Associates Press, New Jersey

44. Cooper, R. and Kaplan, R.S. (1988) Measure costs right: make the right decision, *Harvard Business Review*, 66, 5, 96-103

45. Cronbach, L.J. (1951) Coefficient alpha and the internal structure of tests, *Psychometrika*, 16, 3, 297-334

46. Cohen, J. (1960) A coefficient of agreement for nominal scales, *Educational and Psychological Measurement*, 20, 37-46

47. Crosbie, P. and Bohn, J. (2005) *Modelling Default Risk: Modelling Methodology*, Moody's KMV Company Press, New York

48. Dambolena, I.G. and Khoury, S.J. (1980) Ratio stability and corporate failure, *Journal of Finance*, 35, 4, 1017-1026

49. Davids, M. (1999) How to avoid the 10 biggest mistakes in CRM, *Journal of Business Strategy*, 20, 6, 22-26

50. Dawson, J. (1995) Strategies for retailers in Europe by 2000, *Seminar on Facing Up to The Challenges in Retailing*, Rome

51. Dawson, J. (2000) Retailing at century end: Some challenges for management and research, *International Review of Retail, Distribution and Consumer Research*, 10, 2, 119-148

52. Dawson, J. (2001_a) Innovation, new commerce and European retailing: The next five years, *European Retail Digest*, 30, 13-17

53. Dawson, J. (2001_b) Is there a new commerce in Europe? *International Review of Retail, Distribution & Consumer Research*, 11, 3, 287-299

54. Dawson, J. and Larke, R. (2004) Japanese retailing through the 1990s: retailer performance in a decade of slow growth, *British Journal of Management*, 15, 1, 73-94

55. Deakin, E.B. (1972) A discriminant analysis of predictors of business failure, *Journal of Accounting Research*, 10, 1, 167-179

56. Deakin, E.B. (1976) Distributions of financial accounting ratios: some empirical evidence, *Accounting Review*, 51, 1, 90-96

57. Denscombe, M. (1998) *The Good Research Guide : For Small-Scale Social Research Projects*, Open University Press, Buckingham, UK

58. Dickson, P.R. (1996) The static and dynamic mechanics of competition: A comment on hunt and morgan's comparative advantage theory, *Journal of Marketing*, 60 ,4, 102-106

59. Dietrich, J. R. (1984) Discussion of methodological issues related to the estimation of financial distress prediction models, *Journal of Accounting Research*, 22 Supplement, 83-86

60. Dimitras, A.L., Slowinski, R., Sumaga, R. and Zopounidis, C. (1999) Business failure prediction using rough sets, *European Journal of Operational Research*, 114, 2, 263-280

61. Edmister, R.O. (1972) An empirical test of financial ratio analysis for small business failure prediction, *Journal of Financial and Quantitative Analysis*, 7, 2, 1477-1493

62. Eisenbeis, R.A. (1977) Pitfalls in the application of discriminant analysis in business, finance and economics, *Journal of Finance*, 32, 3, 875-900

63. Eisenbeis, R.A. (2004) *Problems in Applying Discriminant Analysis in Credit-Scoring Models*, in Thomas, L.C., Edelman, D.B. and Crook, J.N. (2004) *Readings in Credit Scoring : Foundations, Developments, and Aims*, Oxford University Press, Oxford, England

64. Elam, R. (1975) The effect of lease data on the predictive ability of financial ratios, *Accounting Review*, 50, 1, 25-42

65. Elmer, P.J. and Borowski, D.M. (1988) An expert system approach to financial analysis: The case of S&L bankruptcy, *Financial Management*, 17, 3, 66-76

66. Fan, A. and Palaniswami, M. (2000) Selecting bankruptcy predictors using a support vector machine approach, Paper presented at *2000 IEEE-INNS-ENNS International Joint Conference on Neural Networks*, 6, 354-359.

67. Fernandez, E. and Olmeda, I. (1995) Bankruptcy prediction with artificial neural networks, *Lecture Notes in Computer Science*, 930, 1142-1146

68. Fielding, J.L. and Gilbert, G.N. (2000) *Understanding Social Statistics*, Sage Publications Ltd., London

69. Fisher, R.A. (1936) The use of multiple measurements in taxonomic problems, *Annals of Eugenics*, 7, 179-188

70. Fitch Ratings (2000) *Assigning Credit Ratings to European Retailers*, Fitch Ratings Press, New York (Downloadable from website: http://www.fitchratings.com/)

71. Fitch Ratings (2001) *Corporate: Corporate Rating Methodology*, Fitch Ratings Press, New York. (Downloadable from website http://www.fitchratings.com/)

72. Foster, G. (1978), *Financial Statement Analysis*, Prentice-Hall Press, New York

73. Franklin, C.M. (1981) Small business observations, *American Journal of Small Business*, 6, 2, 1-2

74. Frydman, H., Altman, E.I. and Kao, D.L. (1985) Introducing recursive partitioning for financial classification: the case of financial distress, *Journal of Finance*, 40, 1, 269-291

75. Garland, R. (2002) Non-financial drivers of customer profitability in personal retail banking, *Journal of Targeting, Measurement and Analysis for Marketing*, 10, 3, 233-248

76. Gentry, J.A., Newbold, P. and Whitford, D.T. (1985) Classifying bankrupt firms with funds flow components, *Journal of Accounting Research*, 23, 1, 146-160

77. Gibson, C.H. and Frishkoff, P.A. (1983), *Financial Statement Analysis: Using Financial Accounting Information*, Wadsworth Inc Press, USA

78. Gilbert, L.R., Menon, K. and Schwartz, K. B. (1990) Predicting bankruptcy for firms in financial distress, *Journal of Business, Finance and Accounting*, 17, 1, 161-171

79. Gilbert, N. (2001) *Researching Social Life*, Sage Publications Ltd., London

80. Gluck, M.A, and Myers, C.E. (2001) *Gateway to Memory: An Introduction to Neural Network Modelling of the Hippocampus and Learning*, MIT Press, USA

81. Gombola, M.J. and Ketz, J.E. (1983) A note on cash flow and classification patterns of financial ratios, *Accounting Review*, 58, 1,105-114

82. Gombola, M.J., Haskins Jr., M.E., Ketz, E. and Williams, D.D. (1987) Cash flow in bankruptcy prediction, *Financial Management*, 16, 4,55-65

83. Grant, R.M. (1991) The resource-based theory of competitive advantage: implications for strategy formulation, *California Management Review*, 33, 3, 114-133

84. Grice, J.S. and Ingram, R.W. (2001) Tests of the generalizability of Altman's bankruptcy prediction model, *Journal of Business Research*, 54, 1, 53-61

85. Gronroos, C. (1996) Relationship marketing: Strategic and tactical implications, *Management Decision*, 34, 3, 5-14

86. Gupta, M.C. and Huefner, R.J. (1972) A cluster analysis study of financial ratios and industry characteristics, *Journal of Accounting Research*, 10, 1, 77-95

87. Gupta, Y.P., Rao, R.P. and Bagchi, P.K. (1990) Linear goal programming as an alternative to multivariate discriminant analysis: A note, *Journal of Business, Finance and Accounting*, 17, 4, 593-598

88. Hair Jr., J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. (1998) *Multivariate Data Analysis*, Prentice-Hall International Inc. Press, USA

89. Hamer, M.M. (1983) Failure prediction: Sensitivity of classification accuracy to alternative statistical methods and variable sets, *Journal of Accounting and Public Policy*, 2, 4, 289-307.

90. Hand, D.J. (1997) *Construction and Assessment of Classification Rules*, John Wiley & Sons Ltd, Chichester, England

91. Harrison, J.S. (2003) *Strategic Management of Resources and Relationships : Concepts and Cases*, Wiley Press, New Jersey

92. Hasty, R. and Reardon, J. (1997) *Retail Management*, McGrew-Hill Companies Inc., New York

93. Hearst, M.A. (1998) Trends and Controversies: Support Vector Machines, *IEEE Intelligent Systems*, 13, 1, 18-28

94. Hjelt, P. (2002) Global 500 Companies, *Fortune Magazine*, 16 August, 54-60

95. Hosmer, D.W. and Lemeshow, S. (2000) *Applied Logistic Regression*, John Wiley & Sons Press, USA

96. Hudson, J. (1986) An analysis of company liquidations, *Applied Economics*, 18, 2, 219-235

97. Hughes II, K.E. (2000) The value relevance of nonfinancial measures of air pollution in the electric utility industry, *Accounting Review*, 75, 2, 209-228

98. Hunt, S.D. (1997) Competing through relationships: grounding relationship marketing in resource-advantage theory, *Journal of Marketing Management*, 13, 5, 431-445

99. Hunt, S.D. (2000) *A General Theory of Competition: Resources, Competences, Productivity, Economic Growth*, Sage Publications Inc., California

100. Hunt, S.D. and Arnett, D.B. (2001) Competition as an evolutionary process and antitrust policy, *Journal of Public Policy and Marketing*, 20, 1, 15-26

101. Hunt, S.D. and Arnett, D.B. (2003) Resource-advantage theory and embeddedness: Explaining R-A theory's explanatory success, *Journal of Marketing, Theory and Practice*, 11, 1, 1-17

102. Hunt, S.D. and Morgan, R.M. (1997) Resource-advantage theory: A snake swallowing its tail or a general theory of competition? *Journal of Marketing*, 61 4, 74-82

103. Ittner, C.D. and Larcker, D.F. (1998) Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction, *Journal of Accounting Research*, Supplement, 36, 3, 1-35

104. Ittner, C.D., Larcker, D.F. and Randall, T. (2003) Performance implications of strategic performance measurement in financial services firms, *Accounting, Organizations and Society*, 28, 7/8, 715-741

105. Johnson, H.T. and Kaplan, R.S. (1987) *Relevance Lost: The Rise and Fall of Management Accounting,* Harvard Business School Press, Boston

106. Jones, E.P., Mason, S.P. and Rosenfeld, E. (1984) Contingent claims analysis of corporate capital structures: an empirical investigation, *Journal of Finance*, 39, 3, 611-627

107. Jones, F.L. (1987) Current techniques in bankruptcy prediction, *Journal of Accounting Literature*, 6, 131-164

108. Kanji, G.K. (1999) *100 Statistical Tests*, Sage Publications, London

109. Kaplan, R.S. and Norton, D.P. (1992) The balanced scorecard—Measures that drive performance, *Harvard Business Review*, 70, 1, 71-79

110. Kaplan, R.S. and Norton, D.P. (1996) *The Balanced Scorecard: Translating Strategy into Action*, Harvard Business School Press, Boston

111. Kaplan, R.S. and Norton, D.P. (2001) *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment,* Harvard Business School Press, Boston

112. Keasey, K. and McGuinness, P. (1990) The failure of UK industrial firms for the period 1976-1984, logistic analysis and entropy measures, *Journal of Business, Finance and Accounting*, 17, 1, 118-135

113. Keasey, K. and Watson, R. (1987) Non-financial symptoms and the prediction of small company failure, *Journal of Business, Finance and Accounting*, 14, 3, 335-354

114. Keasey, K. and Watson, R. (1991) Financial distress prediction models: A review of their usefulness, *British Journal of Management*, 2, 2, 89-102

115. Ketz, J.E. (1978) The effect of general price-level adjustments on the predictive ability of financial ratios, *Journal of Accounting Research*, Supplement, 16, 273-284

116. Knott, A.M. and Posen, H.E. (2005) Is failure good? *Strategic Management Journal*, 26, 7, 617-641

117. Koh, H.C. (1992) The sensitivity of optimal cutoff points to misclassification costs of type I and type II errors in the going-concern prediction context, *Journal of Business, Finance and Accounting*, 19, 2, 187-197

118. Kotha, S. (1995) Mass customization: Implementing the emerging paradigm for competitive advantages, *Strategic Management Journal*, 16, 5, 21-42

119. Kuo, H.C., Wang, L.H., Sheu, H.J. and Li, F.K. (2003) Credit evaluation for small and medium-sized enterprises by the examination of firm-specific financial ratios and non-financial variables: evidence from Taiwan, *Review of Pacific Basin Financial Markets and Policies*, 6, 1, 5-20

120. Kvale, S. (1996) *Interviews: An Introduction to Qualitative Research Interviews*, SAGE Publications Inc., California

121. Lachenbruch, P. (1975) *Discriminant Analysis*, Hafner Press, New York

122. Laitinen, T. and Kankaanpaa, M. (1999) Comparative analysis of failure prediction methods: the Finnish case, *European Accounting Review*, 8, 1, 67-92

123. Lane, W.R., Looney, S.W. and Wansley, J.W. (1986) An application of the COX proportional hazards model to bank failure, *Journal of Banking and Finance*, 10, 4, 511-531

124. Largay III, J.A. and Stickney, C.P. (1980) Cash flows, ratio analysis and the W.T. Grant Company bankruptcy, *Financial Analysts Journal*, 36, 4, 51-54

125. Lee, Y.J. (2001) *Support Vector Machines in Data Mining*, Doctor of Philosophy Dissertation, University of Wisconsin

126. Lee, Y.J., Mangasarian, O.L. and Wolberg, W.H. (2000) Breast cancer survival and chemotherapy: a support vector machine analysis, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, American Mathematical Society, 55, 1-10

127. Lensberg, T., Eilifsen, A. and McKee, T.E. (2006) Bankruptcy theory development and classification via genetic programming, *European Journal of Operational Research*, 169, 2, 677-697

128. Lev, B. (1974), *Financial Statement Analysis : A New Approach*, Prentice-Hall Press, USA

129. Libby, R. (1975) Accounting ratios and the prediction of failure: some behavioral evidence, *Journal of Accounting Research*, 13, 1, 150-161

130. Lincoln, M. (1982) *An Empirical Study of the Usefulness of Accounting Ratios to Describe Levels of Insolvency Risk*, Doctor of Philosophy Dissertation, University of Melbourne

131. Liu, J. (2004) Macroeconomic determinants of corporate failures: evidence from the UK, *Applied Economics*, 36, 9, 939-945

132. Longstaff, F.A. and Schwartz, E.S. (1995) A simple approach to valuing risky fixed and floating rate debt, *Journal of Finance*, 50, 3, 789-819

133. Luoma, M. and Laitinen, E. (1991) Survival analysis as a tool for company failure prediction, *Omega*, 19, 6, 673-678

134. Marchington, M. (1994) *Adding value through the human resource function: The case of food retailin.* in McGoldrick (1994) *Cases in Retail Management*, Pitman Publishing, London

135. Mar-Molinero, C. and Serrano-Cinca, C. (2001) Bank failure: a multidimensional scaling approach, *European Journal of Finance*, 7, 2, 165-183

136. Marais, M.L., Patell, J.M. and Wolfson, M.A. (1984) The experimental design of classification models: An application of recursive partitioning and bootstrapping to commercial bank loan classifications, *Journal of Accounting Research*, 22, Supplement, 87-114

137. McGoldrick, P.J. (2002) *Retail Marketing*, McGraw-Hill Education, New York

138. McGurr, P.T. and DeVaney, S.A. (1998) Predicting business failure of retail firms: An analysis using mixed industry models, *Journal of Business Research*, 43, 3, 169-176

139. McKee, T.E. (2003) Rough sets bankruptcy prediction models versus auditor signalling rates, *Journal of Forecasting*, 22, 8, 569-586

140. McKee, T.E. and Lensberg, T. (2002) Genetic programming and rough sets: A hybrid approach to bankruptcy classification, *European Journal of Operational Research*, 138, 2, 436-451

141. McKee, T.E. and Greenstein, M. (2000) Predicting bankruptcy using recursive partitioning and a realistically proportioned data set, *Journal of Forecasting*, 19, 3, 219-230

142. Mensah, Y.M. (1983) The differential bankruptcy predictive ability of specific price level adjustments: some empirical evidence, *Accounting Review*, 58, 2, 228-246

143. Mensah, Y.M. (1984) An examination of the stationarity of multivariate bankruptcy prediction models: A methodological study, *Journal of Accounting Research*, 22, 1, 380-395

144. Merrilees, B. and Miller, D. (1996) *Retailing Management: A Best Practice Approach*, RMIT Press, Australia

145. Merton, R.C. (1974) On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance*, 29, 2, 449-470

146. Moody's Investors Service (1998) *Rating Methodology: Industrial Company Rating Methodology*, Moody's Investors Service Inc. Press, USA (Downloadable from website http://www.moodys.com/)

147. Moody's Investors Service Inc. (1999), *Moody's Evaluates Key Credit Issues in Retailing*, Moody's Investors Service Inc. Press, USA (Downloadable from website http://www.moodys.com/)

148. Moody's Investors Service (2002) *Retail Rating Methodology: Moody's Approach to Assessing Key Credit Issues in Retailing*, Moody's Investors Service Inc. Press, USA (Downloadable from website http://www.moodys.com/)

149. Moore, A.W. (2001) Cross-validation for detecting and preventing overfitting, Tutorial Slides, Carnegie Mellon University (Downloadable from website: http://www.cs.cmu.edu/~awm/381/2004/crossvalidation.pdf)

150. Moyer, R.C. (1977) Forecasting financial failure: A re-examination, *Financial Management*, 6, 1, 11-17

151. Newman, A.J. and Cullen, P. (2002) *Retailing: Environment and Operations*, TJ International Press, UK

152. Nihon Keizai Shimbun Inc (2002), *Japan Well Performance Companies Survey*, (Online Available at: http://www.nikkei.co.jp/report/02casma1.html)

153. Nørreklit, H. (2000) The balance on the balanced scorecard—a critical analysis of some of its assumptions, *Management Accounting Research*, 11, 1, 65-88

154. Norton, C.L. and Smith, R.E. (1979) A comparison of general price level and historical cost financial statements in the prediction of bankruptcy, *Accounting Review*, 54, 1, 72-87

155. Ohlson, J.A. (1980) Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research*, 18, 1, 109-131

156. Osuna, E., Freund, R. and Girosi, F. (1997) Training support vector machines: An application to face detection, Presented in *IEEE Conference on Computer Vision and Pattern Recognition*, Puerto Rico.

157. Pantalone, C.C. and Platt, M.B. (1987) Predicting failure of savings and loan associations, *American Real Estate and Urban Economics Association Journal*, 15, 2, 46-64

158. Patton, M.Q. (1990) *Qualitative Evaluation and Research Methods*, Sage Publications, London

159. Peel, M.J., Peel, D.A. and Pope, P.F. (1986) Predicting corporate failure—Some results for the UK corporate sector, *Omega*, 14, 1, 5-12

160. Piesse, J. and Wood, D. (1992) Issues in assessing MDA models of corporate failure: a research note, *British Accounting Review*, 24, 33-42

161. Pinches, G.E., Mingo, K.A. and Caruthers, J.K. (1973) The stability of financial patterns in industrial organizations, *Journal of Finance*, 28, 2, 389-396

162. Pine II, B.J., Victor, B. and Boynton, A.C. (1993) Making mass customization work, *Harvard Business Review*, 71, 5, 108-119

163. Platt, J.C. (1999) *Fast Training of Support Vector Machines Using Sequential Minimal Optimization*, In Schölkopf, B., Burges, C.J.C., Smola, A.J. (1999) *Advances in Kernel Methods: Support Vector Machines*, MIT press, USA

164. Platt, H.D. and Platt, M.B. (1990) Development of a class of stable predictive variables, *Journal of Business, Finance and Accounting*, 17, 1, 31-51

165. Qureshi, A.A., Shim, J.K. and Siegel, J.G. (1998) Artificial intelligence in accounting and business, *National Public Accountant*, 43, 7, 13-18

166. Reinartz, W. and Kumar, V. (2002) The mismanagement of customer loyalty, *Harvard Business Review*, 80, 7, 86-94

167. Rose, P.S., Andrews, W.T. and Giroux, G.A. (1982) Predicting business failure: A macroeconomic perspective, *Journal of Accounting, Auditing and Finance*, 6, 1, 20-31

168. Roslender, R. and Hart, S.J. (2003) In search of strategic management accounting: theoretical and field study perspectives, *Management Accounting Research*, 14, 3, 255-279

169. Ross, S.A., Westerfield, R.W. and Jaffe, J. (1999) *Corporate Finance*, Irwin/McGraw-Hill Press, New York

170. Rudolph, T.A., Busch, A. and Busch, S. (2000) Retail food failure and recovery strategies in Switzerland, *Journal of Marketing Channels*, 7, 3, 69-91

171. Said, A.A.; HassabElnaby, H.R. and Wier, B. (2003) An empirical investigation of the performance consequences of nonfinancial measures, *Journal of Management Accounting Research*, 15, 193-223

172. Salchenberger, L.M., Mine Cinar, E. and Lash, N.A. (1992) Neural networks: A new tool for predicting thrift failures, *Decision Sciences*, 23, 4, 899-916

173. Sanchez, R. and Heene, A. (1997) Reinventing strategic management: New theory and practice for competence-based competition, *European Management Journal*, 15, 3, 303-317

174. Sanger, G.C. and McConnell, J.J. (1986) Stock exchange listings, firm value, and security market efficiency: The impact of NASDAQ, *Journal of Financial and Quantitative Analysis*, 21, 1, 1-25.

175. Schölkopf, B., Burges, C.J.C., Smola, A.J. (1999) *Advances in Kernel Methods: Support Vector Machines*, MIT press, USA

176. Sharma, S. and Mahajan, V. (1980) Early warning indicators of business failure, *Journal of Marketing*, 44, 4, 80-89

177. Shank, J.K. and Govindarajan, V. (1992) Strategic cost management: The value chain perspective, *Journal of Management Accounting Research*, 4, 179-197

178. Sheth, J.N. and Parvatiyar, A. (1995) Relationship marketing in consumer markets: antecedents and consequences, *Journal of the Academy of Marketing Science*, 23, 4, 255-271

179. Shimko, D., Tejima, H. and van Deventer, D. (1993) The pricing of risky debt when interest rates are stochastic, *Journal of Fixed Income*, September, 58-66

180. Simmonds, K. (1981) Strategic management accounting, *Management Accounting*, 59, 4, 26-29

181. Smith, M. (1997) *Strategy Management Accounting: Text and Cases*, Reed International Books Pty Limited, Oxford, England

182. Smith, M. (2005) *Performance Measurement and Management : A Strategic Approach to Management Accounting*, Sage Publications Ltd., London

183. Smith, N.C. and Cooper-Martin, E. (1997) Ethics and target marketing: The role of product harm and consumer vulnerability, *Journal of Marketing*, 61, 3, 1-20

184. Srinivasan, V. and Kim, Y.H. (1988) Designing expert financial systems: A case study of corporate credit management, *Financial Management*, 17, 3, 32-44

185. Srinivasan, V. and Ruparel, B. (1990) CGX: An expert support system for credit granting, *European Journal of Operational Research*, 45, 2/3, 293-308

186. Standard and Poor's (2002) *Standard and Poor's 2002 Corporate Rating Criteria*, The McGraw-Hill Companies press, USA (Downloadable from website: http://www.standardandpoors.com/)

187. Standard and Poor's (2003) *Standard and Poor's 2003 Corporate Rating Criteria*, The McGraw-Hill Companies press, USA (Downloadable from website: http://www.standardandpoors.com/)

188. Standard and Poor's (2005) *Sovereign Risk Indicators: Glossary of Terms*, The McGraw-Hill Companies press, USA (Downloadable from website: http://www.standardandpoors.com/)

189. Steltner, H., Staats, R., Timmer, J., Vogel, M., Guttmann, J., Matthys, H., and Virchow, J.C. (2002) Diagnosis of sleep apnea by automatic analysis of nasal pressure and forced oscillation impedance, *American Journal of Respiratory and Critical Care Medicine*, On-line Data Supplement, 165, 7, 940-944 (Downloadable from website: http://webber.physik.uni-freiburg.de/~jeti/)

190. Sung, K.T., Chang, N. and Lee, G. (1999) Dynamics of modeling in data mining: Interpretive approach to bankruptcy prediction, *Journal of Management Information Systems*, 16, 1, 63-85

191. Tabachnick, B.G. and Fidel, L.S. (2000) *Using Multivariate Statistics*, Allyn and Bacon Press, Boston

192. Taffler, R.J. (1982) Forecasting company failure in the UK using discriminant analysis and financial ratio data, *Journal of the Royal Statistical Society*, Series A, 145, 3, 342-358

193. Taffler, R.J. (1983) The assessment of company solvency and performance using a statistical model, *Accounting and Business Research*, 13, 52, 295-307

194. Taffler, R.J. (1984) Empirical models for the monitoring of UK corporations, *Journal of Banking and Finance*, 8, 2, 199-227

195. Takahashi, K., Kurokawa, Y. and Watase, K. (1984) Corporate bankruptcy prediction in Japan, *Journal of Banking and Finance*, 8, 2, 229-247

196. Tam, K.Y. and Kiang, M.Y. (1992) Managerial applications of neural networks: the case of bank failure predictions, *Management Science*, 38, 7, 926-947

197. Tauber, E.M. (1972) Why do people shop? *Journal of Marketing*, 36, 4, 58-61

198. Teece, D.J., Pisano, G. and Shuen, A. (1997) Dynamic capabilities and strategic management, *Strategic Management Journal*, 18, 7, 509-533

199. Tennyson, B.M., Ingram, R.W. and Dugan, M.T. (1990) Assessing the information content of narrative disclosure in explaining bankruptcy, *Journal of Business, Finance and Accounting*, 17, 3, 391-410

200. Thomas, L.C., Edelman, D.B. and Crook, J.N. (2002) *Credit Scoring and Its Applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA.

201. Thomas, L.C., Edelman, D.B. and Crook, J.N. *(2004) Readings in Credit Scoring Foundations, Developments, and Aims,* Oxford University Press, UK

202. TRC (2004), *Rating Criteria Corporate Rating Methodology*, Taiwan Rating Corporation Press, (Online available at: http://www.taiwanratings.com/en/)

203. Trigueiros, D. and Taffler, R. (1996) Neural networks and empirical research in accounting, *Accounting and Business Research*, 26, 4, 347-355

204. Varley, R. and Rafiq, M. (2004) *Principles of Retail Management*, Palgrave Macmillan Press, New York

205. Vedder, R.G. (1987) The opportunities and challenges of managing expert systems, *Industrial Management*, 29, 1, 14-16

206. Yen, Y.C. (1994) *Morden Statistics*, San Min Publishing (In Chinese), Taiwan

207. Walters, D. and Hanrahan, J. (2000) *Retail Strategy: Planning and Control*, Macmillan Press Ltd Press, London

208. Ward, T.J. (1994) Cash flow information and the prediction of financially distressed mining, oil and gas firms: a comparative study, *Journal of Applied Business Research*, 10, 3, 78-86

209. Weber, R. (2004) The rhetoric of positivism versus interpretivism: A personal view, *MIS Quarterly*, 28, 1, iii-xii

210. Wheelen, T.L. and Hunger, J.D. (2003) *Concepts in Strategic Management and Business Policy*, Pearson Prentice Hall Press, USA

211. Wilcox, J.W. (1973) A prediction of business failure using accounting data, *Journal of Accounting Research*, Supplement, 11, 3, 163-179

212. Williams, W.H. and Goodman, M.L. (1971) A statistical grouping of corporations by their financial characteristics, *Journal of Financial and Quantitative Analysis*, 6, 4, 1095-1104

213. Wilson, R. and Sharda, R. (1994) Bankruptcy prediction using neural networks, *Decision Support Systems*, 11, 431-447

214. Wu, C.Y. (2004) Using non-financial information to predict bankruptcy: A study of public companies in Taiwan, *International Journal of Management*, 21, 2, 194-201

215. Wu, W.I. (2005) *Business Research Method*, HWA TAI Publishing (In Chinese), Taiwan

216. Zavgren, C.V. (1983) The prediction of corporate failure: The state of the art, *Journal of Accounting Literature*, 2, 1-37

217. Zavgren, C.V. (1985) Assessing the vulnerability to failure of American industrial firms: A logistic analysis, *Journal of Business, Finance and Accounting*, 12, 1, 19-45

218. Zhang, G.P., Hu, M.Y., Patuwo, B.E. and Indro, D.C. (1999) Artificial neural networks in bankruptcy prediction: general framework and cross-validation analysis, *European Journal of Operational Research*, 116, 1, 16-32

219. Zimmerman, R.M., Kaufman, R.M., Finerty, G.S. and Egan, J.O. (1990) *Retail Accounting and Financial Control*, John Wiley and Sons Press, New York

220. Zmijewski, M.E. (1984) Methodological issues related to the estimation of financial distress prediction models, *Journal of Accounting Research*, 22 Supplement, 59-82

Part Eight

Appendices

- Appendix A: ***Interview Transcriptions***
- Appendix B: ***Pilot Interview Transcriptions and Reflections***
- Appendix C: ***Performance Measures Arrangement***
- Appendix D: ***Performance Measures Regrouping (Based on data Availability)***
- Appendix E: ***E-Questionnaire (English Version)***
- Appendix F: ***E-Questionnaire (Mandarin Version)***
- Appendix G: ***Survey Descriptive Analysis (Mean and Median Data)***
- Appendix H: ***Kolmogorov-Smirnov Test for Department Comparison***
- Appendix I: ***Survey Descriptive Analysis (Case Study)***
- Appendix J: ***Publications***

Appendix A: Interview Transcriptions

Interview (Retail Management - 1)

Interviewee: CFO in a Philippine Chain Store Corporation

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 3 July, 2004 (1400 ~ 1430)

Q1. What are the most important performance measures in the retail industry? Or what are the key factors leading retail company to success?

There are four important performance measures in the retail industry. They are market position, management ability, management ambition and social responsibility.

Q2. Why do you think they are important?

- **Market Position/Target Customers**

There are many different sectors in the retail industry, such as chain stores, supermarkets and department stores. The demand of customers among these market sectors will be different, since each sector has different characteristics, such as product mix. In other words, different retail sectors have different target markets. If a retail company cannot identify its target customers well, it will not have a good performance. Therefore, the ability of a retail company to understand its target customers' needs and satisfactions is a very important performance measurement factor in the retail industry.

- **Management Ability**

With regards to management ability, I will focus on the question as follow: how can we continue our routine work if a staff is expected to vacate his job soon? If a company has a very good internal management system, the impact of this question will be lower. Therefore, the completeness of a retail company's internal management system is also a significant performance measurement factor in the retail industry.

- **Management Ambition**

Every retail company sets its annual operation objectives year by year. However, achieving these objectives depends on management ambition. If a retail company cannot achieve these objectives, it implies that the company cannot run its business well. In other words, the company does not have good performance during these years. However, it is also important to examine whether these objectives are difficult to achieve or not? Because many companies set higher standard for their annual objectives in order to encourage employees to face challenges. In this situation, even if the company cannot achieve its annual objective year by year, it does not mean the company's management does not have high ambition to face challenge.

- **Social Responsibility**

The social responsibility of a retail company becomes more and more important recently. If a retail company has a good image in terms of social responsibility, it will have many added values, such as, increasing the loyalty of customers. Therefore, the factor of social responsibility cannot be ignored in measuring a retail company's performance.

Q3. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

- **<u>Market Position</u>**

How to measure the market position performance of a retail company? It can be achieved by examining the ability of a retail company to satisfy its target customers' needs. Thus, two measures can be considered in this issue:

- **The speed or the frequency of new products development**

This includes two parts: introduction of new products and elimination of dead items (items with low sales). It can be evaluated by the change or turnover rates of the number of new products and the number of dead items. In a normal situation, the two ratios had better be equal to or greater than zero. If the value of the two ratios are negative, this may imply that a retail company decreases its effort to examine the demand of target markets,

  - ((The number of new products introduced at T_1 – The number of new products introduced at T_0)) / The number of new products introduced at T_0) * 100%
  - ((The number of eliminated at T_1 – The number of eliminated at T_0)) / The number of eliminated at T_0) * 100%

- **The customers' demand information collection ability**

The customers' demand information can be collected by POS system (Point-of Sales-system). From this system, we can get information such as, sales structure and customers' structure. We can know the most popular products during a specific time period and who are the major customers, such as young people or old people, who buy these products. The system is very important for a retail company to understand their customers' demand.

- **Comparison analysis with other retail companies**

This comparison analysis usually focuses on many financial and operational ratios. The purpose of this analysis is to try to understand the market competition situation. Moreover, we usually do this comparison analysis not only using the data in a specific year, but also looking at the data during a time period.

- **<u>Management Ability and Management Ambition</u>**

The most important measures in order to assess a retail company's management ability and ambition are as follows:

  - The growth rate of a retail company's sales
  - The growth rate of a retail company's gross profits
  - The growth rate of a retail company's net profits

We usually set these three ratios under a series specific level as annual objectives. Therefore, if a company can achieve these goals in a specific year, this means that the ability and ambition of management has good performance in this year.

- **<u>Social Responsibility</u>**

How to measure the performance of a company's society responsibility? I think the best way is to look at the survey, which is organized by magazines or newspapers, about the ranking of what are the most popular retail companies for new graduates. Maybe there are many different motivations for a new graduate to choose his or her first job; I believe most students prefer companies with good image than companies with bad image. Therefore, you can evaluate a retail company's performance by referring this kind of survey.

Interview (Retail Management - 2)

Interviewee: CFO in a Taiwanese On-line Book Shop

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 7 July, 2004 (1030 ~ 1100)

Q1. What are the most important performance measures in the e-business retail industry? Or what are the key factors leading e-business retail company to success?

Five important factors, which are management ability, strategic vision, organization flexibility, human resource management and crisis management ability.

Q2. Why do you think they are important?

- **<u>Management Ability</u>**

The management ability of a company's leader is very important, especially his or her thoughts. No company can run well without good management ability.

- **<u>Strategic Vision</u>**

Strategic vision is important, since it creates a stable working environment. If a company does not have a clear strategic vision, employees will not know what the right direction of their job is. Employees will become anxious and the performance of their work will reduce.

- **<u>Organization Flexibility</u>**

In an e-business retail company, the organization flexibility is a very significant performance factor, since the technology environment changes very fast. Thus, if an e-business company's organization is not flexible, it is very hard for the company to face the change of technology. The result will be very serious, since in e-business industry, technology can decide a company's performance.

- **<u>Human Resource Management</u>**

Unlike other industries, e-business companies always prefer to hire people with experience. It means that they usually don't hire fresh people, since these people cannot contribute their performance immediately. This will cause a problem that these professional people are very difficult to corporate with other people, since they think they have ability to handle everything alone. Therefore, how to increase the spirits of team work in an e-business company is a very important issue in human resource management. If an e-business company can deal with this problem well, its overall performance will be higher.

- **Crisis Management Ability**

The transactions on the internet have many risks. For example, if an e-business company cannot protect its customers' data well, it will create huge harm to the image of the company. How can a company deal with this problem? It depends on the company's crisis management ability.

Q3. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

All the factors mentioned above are relative to management aspect. How to measure the impacts of these factors, we usually use some ratios as follows:

- **Annual Objectives Achievement Rate**

We usually set series annual objectives for each department. Therefore, the most common measure for evaluating my company's performance is the annual objectives achievement rate.

- **Productivity of Employee**

With regards to productivity of employee, we measure it in terms of two ratios: Sales per Employee and Sales per human resource cost

Actually, sales per human resource cost is more accuracy, since it measures each dollar spent in the human resource's sales productivity.

- **Financial ratios**

Such as current ratio, debt ratio and other financial ratios. Financial ratios express a company's performance at a specific timing or during a time period.

- **Project Perform Rate**

In my company, we usually examine the projects completion progress during a specific time scale. The project perform rate can help us to control the performance and quality of each project and ensure the project progress will not be delay.

Interview (Retail Management - 3)

Interviewee: Financial Planning Manager in a Taiwanese Chain Store Corporation.

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 18 July, 2004 (1530 ~ 1600)

Q1. What are the most important performance measures in the retail industry? Or what are the key factors leading to retail company to success?

Five important performance measures in the retail industry. They are operational performance per store, management ability and ambition, industry development trend, sales and gross profit structure, and cash flow operation.

Q2. Why do you think they are important?

- **Operational Performance per Store**

In order to evaluate a retail company's performance, we first look at this company's store performance. It is very important, since it expresses the performance of a retailer's primary business line. If the stores operational performance of a company is good, it implies that this company is more stable than other companies.

- **Management Ability and Ambition**

Management ability and ambition is also a key factor in retail operation. The most important issue is that the knowledge or know-how of retailer's management is enough to run the business. Moreover, they have to understand the future trend of the industry in order to understand opportunities and threats the company will face in the future.

- **Industry Development Trend**

Every industry has its own business cycle. Different phases of the business cycle will need different business strategies. Therefore, understanding phases of an industry's current and future business cycle is also important to measure performance. Moreover, a retailer's performance is not only based on its internal factors, but also dependent on the external factors. For example, a change of a country's population structure, such as an increase in elderly population, will also have impact on its operation. Thus, industry development trend is also an important consideration in evaluating a company's performance.

- **Sales and Gross Profit Structure**

This issue is very important, since it is relative to a retailer's profitability. For example, if the primary goods of a supermarket are soft drinks, then the profitability of this supermarket is likely to be higher than other competitors, since soft drinks usually have higher gross profits than other products. Therefore, if you want to evaluate a retailer's performance, you cannot omit the examination the company's sales and gross profit structure. This information is available from company's financial statements, such as income statement and annual report.

- **Cash Flow Operation**

Cash flow operation is very important to a company's future development. If a company invests in fixed assets, such as POS system, the investment will enhance a company's future operation. There is an important issue to be considered. If a company intends to do a long-term investment, this company has to consider the synergy with the target company. For example, if a retailer invests in a mining company, there will not be any benefits to either company. On the other hand, if a retailer invests in a distribution centre company, the investment will create synergy and will also enhance these two companies' future operation.

Q3. How do you measure the impact of these factors? What kind of measures you will choose in order to assess this impact?

- **Operational Performance per Store**

The most common measure of operational performance per store is the PSD. (Sales per store per day) Moreover, the most important expenses of a store are rent and human resource costs. In order to measure operational performance per store, these two costs also have to be considered.

- **<u>Management Ability and Ambition</u>**

About this issue, the best way to evaluate a company's management's ability and ambition is to interview the management. Through interviews, you can understand the thoughts and views of management. You can also look at the annual objectives achievement situation to measure this factor. However, don't just focus on a specific timing. You have to examine the achievement situation over a time period, such as 5 years.

- **<u>Industry Development Trend</u>**

How do you understand retail industry's future development trends? There are many sources from where to get the information. For example, government usually has some publications about the development trends of an industry. They are very good reference materials regarding this issue.

Interview (Retail Management - 4)

Interviewee: Previous a UK Hypermarket Manager and Customer Service Director

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 21 July, 2004 (1100 ~ 1200)

Q1. What are the most important performance measures in the retail industry? Or what are the key factors leading retail company to success?

The main performance measures can be divided into two groups: head office performance measures and store performance measures.

Q2. How do you measure the impact of these factors? What kind of measures you will choose in order to assess such impact?

- **<u>Head Office Performance</u>**

In a retail company, there are many departments in the head office, such as finance department, buying department, logistic department and marketing department. Each department has its own performance measures. For example, in terms of the buying department, one of the performance measures is 'where can we find the cheapest or lowest cost goods?' For the logistic department, possible performance measures are 'How good is the relationship with suppliers?' or 'How much costs can we share cost with our suppliers?' Thus, if you want to examine the performance of head office, you have to consider each department's specific performance measures. Since different department has different performance measures.

- **<u>Store Performance</u>**

With regards to the store performance, I think three factors are very important in evaluating a store's performance.

- Wage Management

Wage cost is usually the largest expense in a retail store. For example, there are almost 500 employees working in a UK hypermarket and their wage comprises almost 6.5 percent of total sales. If a retail store cannot control the labour cost well, it will see that as having a very serious impact on its profits. Therefore, if I want to examine a store's performance, I will first examine the performance of wage management. I will also look at the employee structure of the store, such as the percentage of the number of full time staffs and the percentage of the number of part-time staffs, since high percentage of part-time staffs is usually good to reduce wage costs.

- Inventory Management

There is a very important concept of store management called 'shrinkage'. It implies the loss of sales. There are two different types of loss in a store, which are 'known losses' and 'unknown losses'.

- *Known loss*

Known loss usually refers to waste costs. For example, if a store cannot sell out its inventory, these overstocked goods will be thrown away and become a store's loss. This kind of sales loss is called known loss. It can be checked and recorded every day. Known loss is very difficult to control, since it is very difficult to predict the demand of customers. Moreover, there are many outside influences that will also affect a company's inventory management performance, such as, weather conditions. However, known losses can be reduced by a good information system, such as the POS system. Thus, how to reduce known loss is a significant factor of the performance of a store.

- *Unknown loss*

Unknown loss is loss cannot be checked immediately. It can be calculated after stores count their inventory during a period, such as every six months. After a store counts its current inventory, the store can understand how many products they should have sold but did not sell. In other words, these goods have disappeared. There are many possible explanations of this situation. The most common one is that these products were stolen. Unknown loss can be reduced by a good in-store security system. A good store manager should control unknown loss well, since such unknown loss should not have occurred. In my previous company, it is the most important performance measure of a store manager. If a store's unknown loss is very high, this store manager may lose his or her job. Thus, inventory management is important, since it can ensure sales will not be eroded.

- Sales Growth

There are many methods that can be used to increase a store's sales. For example, changing the product mix in order to satisfy the demand of local people. In Scotland, supermarket usually sells food from Scotland, for example, Scotland milk. Local people usually like to buy these products, since these products are closer to their life.

Interview (Retail Management - 5)

Interviewee: Finance Manager in a Taiwanese Chain Store Corporation

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 21 July, 2004 (1000 ~ 1130)

Q1. What are the most important performance measures in the retail industry? Or what are the key factors leading retail company to success?

I think there are four important factors, which are the growth situation of the primary business line, soundness of corporate internal regulations, ability of product development and ability of technological support.

Q2. How do you measure the impacts of these factors? What kind of measures will you choose in order to assess these impacts?

- **<u>The Growth Situation of Primary Business Line</u>**

A company's annual net profits usually come from two different sources: operational profits (primary business line business line) and non-operational profits (such as, interest income, rent income and investment income). If you want to evaluate a retail company's performance, you first need to understand the situation of the company's operational income. Because the operational income comes from the company's primary business line. Thus, the growth situation of primary business line is very important to evaluate a company's performance.

- **<u>Complete Corporation Internal Regulations</u>**

Every company needs complete and good internal regulations in order to maintain the quality of business administration. Therefore, I think sound corporate internal regulations are also a performance measure of a retail company.

- **<u>Product Development Ability</u>**

Most retail products have a characteristic—they have a short product cycle, since consumer tastes change very fast. If a company has good ability to develop new products in order to satisfy customers' demand, it will have better performance. How can a retail company's product development ability be measured? There are many measures that can be used:

- *The amount of new products introduced in a time period*

In order to evaluate this measure, we usually compare the amount of new products introduced with different companies on the same time scale.

- *The life of new products*

If a new product has long life, it implies that the product development ability of a retail company is good. Why? It means that the company has the ability to satisfy customer's demand well, since the new product can sell for a long time. Thus, the longer the life of a new product, the better the new product development ability of a retail company is.

- **<u>Technology Support Ability</u>**

In the retail industry, since the demand of customers changes very fast, the ability to collect the information on customers' demand is very important. These data can be collected by a good technology system, such as POS system. From this system, a company can obtain information about goods and customers. For example, they can understand what are the most popular goods in summer? How many young people like a new introduced product? ...etc. Besides, the management support system in a retail company is also very important. For example, since there are many different products in a store, a retail company usually needs a good supply chain management system. Moreover, a strong accounting system is also very significant in the retail industry. A retail company usually have many stores. If this company wants to process accounting data, it has to collect accounting data from each store first then combine them as whole company data. Therefore good data collection and data process systems are very important to a retail company.

Interview (Retail Management - 6)

Interviewee: CFO in a Taiwanese Pharmacy Store Company

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 28 July, 2004 (0930 ~ 1000)

Q1. What are the most important performance measures in the retail industry? Or what are the key factors leading retail company to success?

There are two aspects about this issue: a quantitative aspect and a qualitative aspect. For the quantitative aspect, we concentrate on four main considerations: sales creation, profits creation, expense control and inventory management. With regards to the qualitative aspect, we usually focus on several important factors, such as, company regulation completeness, human resource education training, differentiate operation, logistic management, marketing management.

Q2. Why do you think they are important? How do you measure these factors?

- **<u>Quantitative Aspect</u>**

- **Sales Creation**

We usually use PSD (Sales per day per store) to evaluate sales in my company. One of the most important objectives in my company is to increase company's PSD. How can we do that? We usually try to increase two other measures: footfall (customer visits) and average transaction size per visit, since an increase of these two measures will increase a company's PSD.

Another question is how can we increase a company's footfall (customer visits) and average transaction size? First, we sell some products that are accepted by most consumers. We call these products NB products (National Brand). Since NB products have very good image in the customers' mind, NB products can increase customer visits and average transaction size. Moreover, we also try to create other sale channels, such as internet (e-business) in order to increase our company's sales.

- **Profit Creation**

Another important objective in my company is to increase company's profits. How can we achieve this objective? The most important method is to sell products with high gross margin. We call these products PB (Private Brand) products. There are many advantages for us to sell PB products. For example, there is no fair market price of PB products. In other words, we can price PB products in terms of our company's marketing strategy. Moreover, the cost of PB products is lower than other products. Thus, PB products sales can create higher profits in a company. However, we cannot just sell PB products to increase profits, since most PB products are not known by customers. Therefore, we usually try to find a balance between PB and NB products in order to achieve the objectives of sales creation and profits creation.

- **Expense Control**

I think cost down is another method to increasing a company's profits. Therefore, we usually set a series of cost down objectives in order to control my company's expense. The most important expenses in my company are rent expense and human resource expense. I think rent expense is difficult to reduce, since most of them are fixed. However, we can try to reduce human resource expenses and other expenses to increase the profits of my company.

- **Inventory Management**

By checking the inventory turnover situation, we can understand how good the inventory management is in my company. There is a policy in my company in that we do not allow any out of stock situation. Even if there are two or three stocks on the shelf; we still think it is an out of stock situation. Why? Because this kind of situation will reduce a customer's buying intention.

- **<u>Qualitative Aspect</u>**

- **Company Procedures' Completeness**

I think the completeness and consistency of a retail company's procedures are very important, especially for store operation. If there is no consistency in a retail company's store operation regulation, the quality of store service will be lower. Because every customer may face different services from different store staffs.

- **Human Resource Education Training**

Since the retail industry faces its customers directly, good interaction between store staffs and customers is very important. Thus, the most important objective of education training in a retail company is to educate its store staffs to provide high quality service to customers.

- **Differentiate Operation**

Every company has to know its strength. Based on this strength, a company can find its profit point and try to develop its operation strategy. There are many different differentiation methods, such as products and services differentiation, or store layout differentiation. A retail company has to identify its role clearly in order to face serious competition situation. Thus, differentiation is very important in the retail industry.

- **Logistic Management**

A retailer usually has many different suppliers. If a retailer has to face all of its suppliers, the operation is very complex and costs time and money. Therefore, a good distribution centre becomes very important for a retail company to solve this problem. There are many advantages to a good logistic management in the retail industry. The most important advantage is it can simplify store operation and maintain the service quality of a store.

- **Marketing Management**

In my company, we usually have some promotional activities in order to increase sales, such as, special day promotion and seasonal promotion. Moreover, we also try to change store layout in order to inspire customers' buying desire. How to keep good relationship with customers is also very important in the retail industry. Thus, CRM (Customer Relationship Management) development is also important. We develop CRM by issuing customer card… in order to increase the customer loyalty. I think marketing activities are very important in the retail industry, since customers usually change their tastes very fast. Therefore, if I want to evaluate a retail company's performance, I will consider a retail company's market management performance.

Interview (Retail Management – 7)

Interviewee: CFO in a Taiwanese Gas Station Company

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 30 July, 2004 (0930 ~ 1000)

Q1. What are the most important performance measures in the retail industry? Or what are the key factors leading retail company to success?

I think the most important performance measures in the retail industry are: marketing activities, expense control, cash flow operation, human resource training and brand image.

Q2. Why do you think they are important? How do you measure these factors?

- **<u>Marketing Activities</u>**

With regards to marketing management, I think the most important point is differentiation. Why? Because differentiation can inspire the demand of customers to increase sales and profits.

- **<u>Expense Control</u>**

I think good expense control is very important to every company. In my company, I think the most important expenses we need to control are rent expense, buying expense and gift expense. In Taiwan, we usually send gifts to our customers in order to inspire them to our gas station to gas up their car. The gift expenses comprise large part of our total expenses. Expenses will erode a company's profits. Therefore, how to control expenses is very important to a retail company.

- **<u>Cash Flow Operation</u>**

We can discuss about this performance factor in terms of two key points: the ability for a company to find the cheapest funding sources and the ability for a company to use their cash flow. For the first ability, I think a public company will have better performance, since they have more funding sources than private companies, such as funding from the capital market. With regards to the second ability, I think a company with a stable investment strategy will have better performance than other companies that invest their cash flow in risky targets.

- **Human Resource Training**

We hope we can have high quality staffs with professional knowledge and skills. How to achieve this objective? It depends on a good training and education system in a company. Therefore, I think human resource training is also a very important performance factor in the retail industry.

- **Brand Image**

There are many advantages that a company with a good brand image has. For example, good brand image will enhance the customer loyalty and inspire customers to consume. Therefore, if you want to evaluate a retail company's performance, you cannot forget to consider the brand image factor.

Interview (Retail Management – 8)

Interviewee: CFO in US Coffee Shop Company (Taiwan and China)

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 30 July, 2004 (1000 ~ 1030)

Q1. What are the most important performance measures in the retail industry? Or what are the key factors leading retail company to success?

There are five important performance factors in the retail industry: brand strength, business culture, operational procedures, product innovation ability and team work ability.

Q2. Why do you think they are important? How do you measure these factors?

- **Brand Strength**

For my point of view, I think brand strength is the most important performance factor in the retail industry. For example, my company has better brand strength than other coffee companies in Taiwan and China. It is easier for us to create good relationship with our customers, suppliers and government. In other words, we faced less hurdles than other companies, when we start or run our business. Therefore, if a retail company has very strong brand strength, it will perform better than other companies.

- **Business Culture**

In my company, the most important asset is our employees. We usually share all the harvests with our company's staffs and respect them. Therefore, the loyalty of Starbucks staffs is higher than other competitors. I think good business culture is also a key performance factor in the retail industry.

- **Operation Regulation**

Retail is detailed. If you intend to run a retail business, you have to know all the details relative to this business. Moreover, retail is one kind of business related to the 'human being'. You need to make contact with people, since they are your customers. Thus, the retail business is a complex business and you need a complete set of operational procedures in order to maintain your service quality. Complete operational procedures are also important in the retail industry.

- **Product Innovation Ability**

As I mentioned before, retail business is one kind of business related to the 'human being'. Thus, the performance of a retail company depends on the feeling of customers. However, customers' taste or feeling is always changing. How to control customer's demand is also a significant goal in the retail industry. It can be achieved by new product innovation. So, if a retail company has good product innovation ability, it will have better performance. Because the ability of this company to catch customers' demand is better than other companies.

- **Team Work Ability**

I think we do not need a strong person to manage our company, but we need a good team to run our business. Every one has to play his or her role well in the company. We encourage different departments to corporate with each other, since we believe a project with different point of views is more complete. Therefore, the teamwork ability of a retail company is also very important.

Interview (Retail Management – 9)

Interviewee: Finance Manager in a Japanese Department Store Company

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 30 July, 2004 (1100 ~ 1120)

Q1. What are the most important performance measures in the retail industry? Or what are the key factors leading retail company to success?

I think there are four key factors leading retail company to success: management system, marketing strategy, learning organization and financial considerations.

Q2. Why do you think they are important? How do you measure these factors?

- **Management System**

A good management system is very important, especially for a joint venture company. For example, my company is a joint venture company with a Japanese department store company and a Taiwanese Chain Store Corporation. Under this situation, two different parent companies will cause some conflicts in terms of business culture or other factors. How can this problem be solved? How can this joint venture company be run? This depends on a good management system. Moreover, the ability and the vision of management also have a great impact on a retail company's performance. Therefore, good management system includes a complete internal regulation and a strong management team.

- **Marketing Strategy**

I think there are two important strategies related to retail business operation. The first strategy is the differentiation strategy. Every retail company has to know what its market position is and what its strength is. In my company, our market positioning strategy is to give 'surprises' to our customers. Therefore, we try to differentiate our store image and layout in order to achieve our operational objective. The second important strategy is the CRM. (Customer Relationship Management) Since my company is a department store company, how to keep a good relationship with our customers is the most important point to our performance. We issued the key customer card in order to inspire the will of our customers to shop in our store. Thus, I think differentiation strategy and CRM are both very important to our retail operation.

- **Learning Organization**

Another key performance factor is the organization. I think a retail company needs a high-spirited organization, especially a learning organization. Since the change of the customer's demand is very fast in the retail industry, how to transfer a message or communicate in the organization is a significant factor to catch customer's demand. Thus, managers have to have the ability of 'listen'. Moreover, a retail company had better authorize its staffs some power to make some decisions. It will also enhance the efficiency of the operational process. Therefore, a learning organization is also an important performance measure in the retail industry.

- **Financial Considerations**

With regards to the financial aspect, I think there are some measures that are very important. The growth of sales is the first factor to measure in a retail company's performance, especially for the PSD. (Sales per day per store) Second, we will look at a retail company's operating income, since operating income focuses on the profits of a company's primary business line. Finally, we will try to examine a retail company's inventory system, since out of stocks is a very series problem of the retail operation.

Interview (Retail Management – 10)

Interviewee: Finance Manager in a Taiwanese Cleaning Service Store Corporation

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 2 August, 2004 (1030 ~ 1100)

Q1. What are the most important performance measures in the retail industry? Or what are the key factors leading retail company to success?

Five key performance factors in the retail industry: brand strength, resource sharing, management, customer orientation and technology development.

Q2. Why do you think they are important? How do you measure these factors?

- **Brand Strength**

Based on my previous experience, a retail company with a good brand image, especially a foreign brand company, is more likely to succeed than other retail companies. I think the reason is that the brand image of these companies has been accepted by most customers in terms thanks to previous performance. Thus, I think brand strength is one of the important performance measures in the retail industry.

- **Resource Sharing**

You can discuss this issue in terms of many different aspects. For example, how to share resource with your suppliers in order to reduce fixed costs? How to share knowledge with your company's internal staff in order to achieve continuous learning objective? I think the main idea is how to keep a good relationship and how to share resources with a retail company's internal and external stakeholders in order to create synergy. I believe that a retail company who shares its resources with other stakeholders will have better performance.

- **Management**

About management, there is a new strategic management approach called the balanced scorecard. We intend to import this approach into our company. The balanced scorecard is based on four perspectives: the learning and growth perspective, the business process perspective, the customer perspective and the financial perspective. It is a very good management approach we can use to examine our company's performance.

- **Customer Orientation and Technology Development**

How to control customer's demand is very important for the retail business operation, since customer's demand directs a retail company's strategy. A retail company must be customer oriented. How to know and catch the customer's demand? It relies on a good information system, such as POS system. Thus, if a retail company cannot control its target customer's demand, it will not have better performance in the future.

Interview (Retail Management – 11)

Interviewee: Finance Manager in a Taiwanese Auto Parts and Tie Stores

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 5 August, 2004 (0930 ~ 1000)

Q1. What are the most important performance measures in the retail industry? Or what are the key factors leading retail company to success?

I think there are five key performance factors, which are: organizational procedures, channel advantage, technology development, logistic management and human resource quality.

Q2. Why do you think they are important? How do you measure these factors?

- **Organizational Procedures**

If a company is controlled by people, not procedures, there will be many shortcomings. For example, people have different preferences and moods. These emotional factors will have great impact on a company's performance. Thus, a retail company with a complete procedural system is very important. How can I evaluate a retail company's organization procedures? One of the measures is the ISO (International Organization for Standardization) system. If a company has passed the examination of ISO, I think this company's organization procedures are more complete than other companies.

- **Channel Advantage**

If a retail company has more stores than other companies, it will have many advantages that other companies do not have, such as store fixed cost reduction. Moreover, since store number is one of the key measures of market share, a retail company with a large number of stores will have greater bargaining power vis-à-vis its customers and suppliers as compared with companies without a large store number. Thus, channel advantage is also an important measure in the retail industry. How can we enhance our company's channel advantage? We usually expand our store number through a franchise system. The most important advantage comes from speed of expansion. Through a franchise system, our company's human resource cost and rent cost will lessen. However, the success of this strategy depends on a good brand image and a complete management system. No one will want to be a franchisee in a bad performance retail company.

- **Technology Development**

A good system in a retail company will have two main advantages. First, since the retail operation process is very detailed, a good management system will enhance the efficiency of the operation process. Second, a good system can help a retail company to obtain customer's information quickly. Hence, if a retail company has good management system and continues to invest in technological R&D, it will have better performance in the future.

- **Logistic Management**

Another important factor to simplify the retail operating process is a good logistic management system. When I want to evaluate a retail company's performance, I will also consider a retail company's logistic system.

- **Human Resource Quality**

We usually evaluate staff's performance in terms of eight factors, as follows: the ability to achieve work objectives, the ambition to expand the job content, communication skills, the relative professional job skill and knowledge, responsibility, work attitude, team work ability and interaction with customers

I think the most important factor is how good the interaction between our store staffs and customers is. Because the retail industry has to face its customers directly. Therefore, we prefer to hire store staffs with a smiling face and friendly personality.

Interview (Retail Management – 12)

Interviewee: Investor relationship manager in a Taiwanese Chain Store corporation; previous CFO in a Philippine Chain Store Corporation.

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 5 August, 2004 (1500 ~ 1530)

Q1. What are the most important performance measures in the retail industry? Or what are the key factors leading retail company to success?

This question can be illustrated by two aspects: internal factor and external factor. About the internal factor, I will focus on the question as follow: how to catch customer's demand? The objective can be achieved by differentiate strategy, the performance of store and head office staffs. Regarding the external factor, I will focus on the political environment impact.

Q2. Why do you think they are important? How do you measure these factors?

- **Internal Factor: How to catch customer's demand?**

- **Differentiation Strategy**

There are many different strategies for differentiation. I think the most important differentiation strategy is the product differentiation strategy. If all retail companies sell the same products, there will be many shortcomings. For example, price competition will become very serious under this situation. Moreover, every company finds it very difficult to identify its market position. Thus, I think product differentiation strategy is very important. How can I measure how good a retail company's product differentiation strategy is? I think there are two measures to refer to. The first is the amount of private brand products. Private brand products have the character of uniqueness. In other words, you can just find them in a particular company. Thus, if a retail company has more private brand products than other companies, I think this company has better performance than other companies. Second, I will also examine how successful a retail company is in developing its new products. It implies how good a retail company's R&D ability is. If a retail company has very good R&D ability, I think this company has better ability to face customer's demand.

- **The Performance of Store and Head Office Staffs**

I think every staff in a company must have the ability and the will to achieve his or her work objective. If one of the criteria is lacking, a staff will not have good performance. This applies to both store and head office personnel. Therefore, how to maintain and enhance a staff's ability and will to achieve company's objective is very important to a company's performance. It can be enhanced by a good training system. Thus, I will also examine a retail company's education and training system in order to evaluate its performance.

- **External Factor: Political Environment Impact**

Why do I mention this issue? Since my previous position was the CFO in a Philippine chain store corporation, I found out some problems about this issue. For example, Philippine's law has many limitations to the retail industry development. We need to spend 3 months to open a new store. It has very serious impact on our channel development strategy. Moreover, the infrastructure in the Philippines is not very complete. For example, if we want to transport our goods from south Luzon to north Luzon, we usually need one day to transport them. It implies that it is very difficult for us to create our logistic system in Philippine. Therefore, based on my previous experience, political environment impact is also an important performance consideration in the retail industry.

Interview (Retail Management – 13)

Interviewee: Human Resource Department Manager in a Taiwanese Chain Store Corporation; previous CFO in a China Coffee Shop Corporation)

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 6 August, 2004 (1000 ~ 1030)

Q1. What are the most important performance measures in the retail industry? Or what are the key factors leading retail company to success?

Every company's extreme objective is to increase financial performance measures, such as net income and sales. There are three important performance factors in the retail industry that help to achieve the financial objective: team, brand strength and know-how.

Q2. Why do you think they are important? How do you measure these factors?

- **Team**

First, I have to define what the team is. The team includes all the stakeholders in a retail company, such as internal staffs, suppliers and customers. How to integrate all the stakeholders' power in order to create synergy is the key issue of this factor. Moreover, how to maintain good relationship with them is also very important.

How can I measure a retail company's team strength? For example, in terms of internal staff, I can measure team strength by checking the average tenure of the employees. High tenure means high employee loyalty. However, you cannot just depend on this measure. The most important thing is that you have to contact them directly in order to know their ability and attitude. It's more objective.

- **Brand Strength**

Brand is the company's image in the customers' mind. Good brand usually can enhance customer loyalty and can satisfy the needs of customers. Thus, it will increase a retail company's potential sales. In order to evaluate a retail company's brand strength, I think the best way is to refer to some market surveys from some magazines, journals or other secondary materials. You can obtain much valuable information through these market surveys.

- **Know-how**

Know-how is a learning process. Every company has its own know-how, since every company's environment is different. A good know-how can increase the efficiency of business operating or reduce the operating mistakes. Thus, I think know-how is also a very important performance measure in the retail industry.

Interview (Lender – 1)

Interviewee: Business Loan Department Manager in a US Bank

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 21 July, 2004 (0930 ~ 1000)

Q1. Before you make a loan decision to a retail company, what are the most important performance measures you will consider? Or what are the key factors leading retail company to success?

In my bank, we usually evaluate a retail company by using two different performance measures: qualitative performance measures and quantitative performance measures. For the quantitative performance measures, we focus on a retail company's ability to face its future obligations, cash flow management and income statement analysis. For the qualitative performance measures, we concentrate on management ability and vision, market position maintenance, brand image, store management, marketing strategy, product innovation and system strength.

Q2. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

- **<u>Quantitative Performance Measures</u>**

- **The ability of a retail company to face its future obligations**

This measure is the most important consideration for us to make a loan decision. If we decide to lend money to a retail company, this company should have ability to face its future obligations. We can use some financial ratios to evaluate this issue, such as cash ratio and current ratio. Moreover, we also evaluate this ability by examining a retailer's ability to obtain capital sources. If a company has many different sources of cash inflow, such as money from stock market, we will regard it as a positive factor of the ability to face its future obligations.

- **Cash Flow Management**

Retail companies usually have more cash flow than other industries, since they usually receive cash from their customers. Thus, how to use large cash flows become a very important issue for a retail company. If a retail company uses their cash flow in their primary business, we usually regard it as a positive sign. However, if a retail company invests its money to a function, which is not related to its primary business, we will pay more attention on this investment. Because this investment will not create synergy to its business.

- **Income Statement Analysis**

Regarding the income statement analysis, we focus on one key point: how a retail company maintains its sales and profits. This can be examined by looking at the sales and profits growth situation during a time period. Moreover, we will also evaluate a company's ability to control its costs and expenses.

- **<u>Qualitative Performance Measures</u>**

- **Management Ability and Vision**

We usually do a series interviews with a retail company's management before we make a loan decision. Through the interviews, we can understand the ability and vision of this company's management. For example, since the retail market in Taiwan is not too large, we think if a company management have good future vision, they should have international expansion plan. In other words, if a Taiwan retail company does not have international expansion plan, we will think this company's management does not have good future vision. Moreover, we will also examine the possibility of success of a retail company's international expansion plan in order to understand the ability of management. Therefore, we think the management ability and vision is a very important performance measurement in the retail industry.

- **Market Position Maintenance**

With regards to this measure, we focus on a market share analysis. Because most retail companies sell low margin products, sales becomes the key factor to increase profits. Thus, how to maintain or increase sales is also a very important performance measure in the retail industry. How do we measure this? We usually evaluate a retail company's plan for channel development. Stores expansion is the fastest way to increase sales, since it can create economies of scale. Moreover, economies of scale can also reduce fixed costs in each store and increase profits. Therefore, we will also evaluate a retail company's ability to maintain its market position before we make a loan decision.

- **Brand Image**

Brand can be regarded as an intangible asset of a company. It also implies the strength of customer's loyalty. We usually prefer to lend money to a company with good brand image, since we believe this company has been accepted by customers. How strong is a company's brand image? You can obtain valuable information from some secondary materials, such as, magazine, since some magazines conduct a survey about a company's brand image every year. From these surveys, you can rank each retail company in terms of brand image.

- **Store Management**

A retail company's profits come from its stores. Thus, good store management is also a significant performance measure. We evaluate this measure by checking a retail store's product arrangement, inventory management and shopping environment. For example, if a company can keep some products in stock, it implies that the company can control its inventory well. Because the company can understand customers' demand and make sure customers can buy the products they want at any time.

- **Marketing Strategy**

Good marketing strategies can also enhance a retail company's market position. For example, if a retail company has some promotion activities for some specific people or dates, its sales will also increase thanks to these activities. Moreover, these activities will also increase the loyalty of customers.

- **Product Innovation**

Not only can a good retail company understand customer demand, but it can also lead customer's demand. How does a retail company achieve this aim? The objective can be achieved by new product innovation.

- **System Strength**

The technology is very important to a retail company, since technology can help a retail company to collect information and to control operation. Thus, I think system strength is also a significant factor for measuring a retail company's performance.

Interview (Lender – 2)

Interviewee: Business Loan Department Manager in a Taiwanese Bank

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 26 July, 2004 (0900 ~ 0930)

Q1. Before you make a loan decision to a retail company, what are the most important performance measures you consider? Or what are the key factors leading to success for a retail company?

Before we make a loan decision, we usually evaluate a retail company's performance by six factors: location, brand strength, product character, competitiveness, management ability and financial factors.

Q2. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

- **Location**

I think different retail formats need different locations. For example, a department store is usually located in the city centre, but a hypermarket is not. If a department store is located in a low population density area, such as the countryside, we think the department store will not have good performance in the future. Moreover, there are some important issues that should be considered, such as, the convenience of transportation and parking. We will consider all these issues in order to evaluate a retail company's location strength.

- **Brand Strength**

Brand strength is very important in the retail industry. We prefer to lend money to a retail company with a famous brand. Why? I think if a retail company has a good brand image, it implies that its products have been accepted by their customers. Moreover, a successful retail company usually has good brand image. It means that good brand image also implies good performance. Therefore, I think brand image is an important performance measure in the retail industry.

- **Product Character**

We also prefer to lend money to a retail company with popular products. If a retail company's goods are not very popular, this company has high operational risk. Because these companies have higher good unsalable risk. Therefore, before we make a loan decision, we will also consider the product character of a retail company.

- **Management Ability**

Before we make a loan decision to a retail company, we usually have the chance to interview the company's management. Through this interview, we can understand this company's future plan, the vision and ambition of this company's management. Moreover, we can evaluate this company's management ability by checking its past performance.

- **Competitiveness**

I think the most important measure relative to a retail company's competitiveness is the market share by sales. If a retail company's sales are above its competitors, we will prefer to lend money to this company. Moreover, we also prefer to lend money to a retail company with a great number of stores. There are many advantages of a retail company with large store numbers, for instance, sharing fixed costs. Therefore, a competition analysis is also very important to evaluate a retail company's performance.

- **Financial Factors**

We evaluate a retail company's financial performance by many financial measures: gross margin, inventory turnover, debt ratio, current ratio, operational margin and interest cover. With regards to the cash flow operation, we hope retail firms can save their cash in our bank, since we can know the retail company's cash flow operation program. It can provide valuable information for us to make a loan decision.

Interview (Lender – 3)

Interviewee: Business Loan Department Manager in a Taiwanese Bank

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 26 July, 2004 (1000 ~ 1020)

Q1. Before you make a loan decision to a retail company, what are the most important performance measures you consider? Or what are the key factors leading to the success of a retail company?

We will consider 5 factors: target company past credit situation, brand strength, management ability, market share and financial considerations.

Q2. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

- **Past Credit Situation**

Before we make a loan decision, the first important task is to evaluate the target company's past credit performance situation. If the target retail company did not have good credit performance in the past, of course, we will not lend money to this company. We can obtain the relative information through the Joint Credit Information Centre in Taiwan.

- **Brand Strength**

We will lend money to a company with good brand image, since we think this company has good performance in the past. Moreover, good brand image also implies the loyalty of customers. I think brand strength is an important performance measure in the retail industry.

- **Management Ability**

I think every company needs a good management. How can we measure a company's management ability? We usually evaluate a company's management ability by some secondary materials, such as reports from Business Week. Moreover, we can also interview with management in order to understand the management's personality and vision.

- **Market Share**

In order to measure a retail company's market position, we usually make a competition analysis in terms of market share by sales. We prefer to lend money to the retail company with large amount of sales, since we think these kinds of companies are more stable than others.

- **Financial Considerations**

With regards to the financial statement analysis, we usually examine a retail company's financial statement for three years. The most important objective is to measure a retail company's ability to face its future obligations. There are some important measures we will consider, such as, sales, account receivable turnover, account payable turnover, inventory turnover and current ratio. Moreover, we will look at a retail company's store number, since we think it is an important performance measure for evaluating a retail company's scale.

Interview (Lender – 4)

Interviewee: Business Loan Department Manager in a Taiwanese Bank

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 26 July, 2004 (1020 ~ 1040)

Q1. Before you make a loan decision to a retail company, what are the most important performance measures you consider? Or what are the key factors leading to the success of a retail company?

There are three main factors: cash flow operation, management ability and relationship with stakeholder, as well as financial analysis.

Q2. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

- **Cash Flow Operation**

Because most retail companies receive cash, they tend to have more cash flow than other industries. Therefore, how to use their cash flow becomes a very important issue. We prefer to lend money to a company that invests its cash flow in it primary business (e.g. fixed assets), for we will then consider this company to have a stable cash flow operation strategy. In other words, the company's risk is lower.

- **Management Ability and Relationship**

I think good management not only requires good management ability, but it also requires good relationship with stakeholders such as banks, outside investors or government. Good relationship with stakeholders will enhance a retail company's strength, because it can more easily access useful resources as compared with other companies.

- **Financial Analysis**

With regards to the financial analysis, I focus on a company's profitability, capital structure (debt ratio) and inventory management. About inventory management, I concentrate on the goods unsalable risk, especially for the food retailers, since food products are highly perishable. We usually use inventory turnover to evaluate a retail company's goods unsalable risk.

Interview (Lender – 5)

Interviewee: Business Loan Department Manager in a Taiwanese Bank

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 26 July, 2004 (1100 ~ 1120)

Q1. Before you make a loan decision to a retail company, what are the most important performance measures you consider? Or what are the key factors leading to the success of a retail company?

With regards to this issue, I think there are four factors we will consider before we make a loan decision: market position, economies of scale, company background and financial consideration.

Q2. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

- **Market Position**

About this factor, we will focus on the measure of market share in terms of sales. We believe that large sales figures imply greater stability or sustainability.

- **Economies of Scale**

No one will argue that the larger a company's profits, the better the company's performance. With regards to a retail company's profits, we need to look at the operational margin, since operational margin is equal to the profits from a retail company's primary business line. Therefore, if a retail company's operational margin is very large, we will prefer to lend money to this company.

- **Company Stakeholders' Background**

This factor is also important, since a retail company with strong stakeholders' background usually implies there is a strong support behind this company. For example, Uni-President Company is the largest food business group in Taiwan. If Uni-President Company invests a retail company in order to create sales channel, we will lend money to this new retail company. Because this new retail company will obtain the support from the Uni-President Company.

- **Financial Considerations**

- **Current Ratio**

We analyze current ratio in order to understand a retail company's ability to face its short-term obligations.

- **Debt Analysis**

With regards to this issue, we will focus on the account payable analysis, since there are many different suppliers in a retail company. By checking the account payable situation, we can understand whether the retail company has good relationship with its suppliers or not.

- **Gross Margin**

There are two reasons why we want to evaluate a retail company's gross margin. The first is to understand the bargaining power of the retail company vis-à-vis its suppliers. If a retail company has high bargaining power vis-à-vis its suppliers, the products' margin obtained will also be high. The second reason is to understand a retail company's future profitability. Given that one of the characteristics of a retail industry is its low margin, having higher gross margin than other retail companies would imply higher competitiveness in the future. In other words, the retail company in question will have higher profitability in the future.

Interview (Lender – 6)

Interviewee: Business Loan Department Manager in a Taiwanese Bank

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 26 July, 2004 (1130 ~ 1150)

Q1. Before you make a loan decision to a retail company, what are the most important performance measures you will consider? Or what are the key factors leading retail company to success?

I think there are four important factors we will consider: economies of scale, company background, law considerations and financial considerations.

Q2. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

- **Economies of Scale**

We evaluate a retail company's economies in terms of three measures: store numbers, sales and profits--especially sales. I think the gross margins of most retail companies' products are low; therefore sales is a key factor to create profits. If a retail company's sales are higher than other companies, we will prefer to lend money to that company.

- **Company Background**

If a retail company's investors are very strong and can support this retail company, we will also prefer to lend the company money. We think such companies are more stable than other companies without strong investor background.

- **Law Consideration**

According to the bank law in Taiwan, a bank cannot make a loan exceeding 5% of the borrowing company's total equity. Therefore, if we have lent money to a specific company before, we need to check what is our total amount of loan to this company.

- **Financial Considerations**

With regards to the financial analysis, we will focus on three years financial statement analysis. We examine a retail company's financial structure, profitability, turnover ratios and the ability of a retail company to face its future obligations. We also evaluate a retail company's cash flow operation. We will regard a retail company as a good performance company, if it has net positive cash inflow.

Interview (Lender – 7)

Interviewee: Business Loan Department Manager in a Taiwanese Government Bank

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 30 July, 2004 (1000 ~ 1030)

Q1. Before you make a loan decision to a retail company, what are the most important performance measures you will consider? Or what are the key factors leading retail company to success?

We usually consider three factors before we make a loan to a company: debtor previous credit history, the use of the loan and the ability of repay.

Q2. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

- **Debtor Previous Credit History**

The first important consideration for us to make a loan decision is to check debtor previous credit history. We can get such information through Joint Credit Information Centre in Taiwan. If the debtor did not have good credit before, we will not lend money to this company.

- **The Use of the Loan**

The purpose of the use of the loan is also an important consideration for us to make a loan decision. We prefer to lend money to a company, which intends to use the money to invest in fixed asset or other capital expenditures. Because these investments are good for this company's future development. However, if a company intends to use the money for a arbitrage purpose (For example, borrow money at lower interest rate and use the money to repay the loan with high interest rate) or for the risky investments, we will consider these situations carefully.

- **The Ability of Repay**

With regards to this factor, we usually examine a retail company's sales. If a retail company's sales is declining now, we will regard it as a negative situation. Because the situation implies that the ability of the company to repay the loan is dropping now. Moreover, we think public companies have better repay ability than private companies, since their credits are more reliable. How can we evaluate a company's repay ability? We usually evaluate the ability in terms of a company's account receivable turnover, account payable turnover and inventory turnover. In other words, we try to understand a company's cash operation pressure. Finally, We usually ask for mortgage targets or repay insurance or repay promise from the borrower before we make a loan decision. It is also a guarantee of a company's repay ability. Therefore, the ability for a retail company to face its future debts is also an important performance factor.

Interview (Lender – 8)

Interviewee: Business Loan Department Manager in a Taiwanese Bank (London Branch)

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 26 August , 2004 (1100 ~ 1120)

Q1. Before you make a loan decision to a UK retail company, what are the most important performance measures you consider? Or what are the key factors leading to the success of a UK retail company?

There are three important factors we will consider before we make a loan decision:

leadership, financial analysis and the use of the loan.

Q2. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

- **Leadership**

I think it's the most important performance measure. If a company's leader is not reliable or does not have good credit situation in the past, we will not lend this company money. How can we measure this factor? I think there are two methods we can use:

- **Debtor's Previous Credit History**

We can obtain the information of the debtor's previous credit history from a credit evaluation institute in the UK. I think it's one kind of secondary data collection method. From this information, we can have a basic idea about the debtor's past credit situation.

- **Market Information**

I think market information is more important than the debtor's previous credit history, since market information is richer than other secondary information. If a debtor has a good credit situation in the past, but its market reputation is not very high, we will reconsider the loan decision. We can get market information through our relationships with many market stakeholders, such as other bankers. Therefore, we usually join many activities in order to understand the market situation.

- **Financial Situation**

- **Financial Structure**

Can a company face its future obligations? I think that is the most important consideration for a banker to make a loan decision. They usually evaluate a company's ability to face its future obligations in terms of some financial ratios, such as, interest cover, debt ratio, current ratio and cash flow sustainability.

- **Profitability**

With regards to the profitability of a company, bankers usually evaluate it in terms of two aspects: current profitability situation and the impact on profitability from a company's future plans. About the current profitability situation, bankers evaluate it by examining a company's financial statements. They usually focus on some important profitability financial ratios, such as ROA, ROE and gross profits. Actually, they pay more attention on a company's future plans impacts. For example, investment in China is a very hot issue these years, since China has a very huge market and low labour costs. If a retail company intends to invest China with a very good and stable plan, the profitability of this company should be improved in the future. Thus, they usually think it as a positive factor of a company's profitability.

- **The Use of the Loan**

Understanding a company's cash operation plan is also a key performance measure for us to make a loan decision. We evaluate a company's performance not only according to its operations plan, but also to its real situation. For example, a retail company needs money to open a new store. In the operations plan, they think they can earn a lot of money from the store in the future. However, we will not lend it money just based on the plan. We will go to see how good the location is and evaluate the feasibility of the plan. Thus, to evaluate the use of the loan is also an important consideration for us to make a loan decision.

Interview (Investor – 1)

Interviewee: Industrial analyst in a US Global Investment Group

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 19 July, 2004 (1200 ~ 1215)

Q1. What are the most important factors relative to your investment decision to a retail company? Why do you think they are important?

It can be divided into two different aspects: operational aspect and financial aspect. Each aspect has its own key considerations. With regards to the operational aspect, we will focus on the company's competition ability and future development trends. For the financial aspect, we will concentrate on the basic financial analysis and comparison analysis with its competitors. Both of these two aspects are important, you cannot just rely on only one of them.

Q2. How do you measure the impact of these factors? What kind of measures you will choose in order to assess this impact?

- **Operational Aspect**

- Competition Ability

With regards to the competition ability of a company, we usually focus on three measures: store number, number of customers per store per day and average transaction size. Normally, the higher these measures are, the better.

- Future Development Trends

Regarding the future development trend of a retail company, we will concentrate on the company's future market share situation in terms of sales or profits.

- **Financial Aspect**

- Basic Financial Analysis

We look at some important market measures to evaluate a company's value, such as P/E ratio. We think that based on an investor point of view, the lower the P/E ratio is, the higher the potential future return. However, a company with a higher P/E ratio also implies better growth power. For example, Hi-tech industry usually have higher P/E ratio than traditional food manufactures. The choice of the P/E ratio depends on the purpose of investment and it varies in terms of different industries. Moreover, we will also examine the product mix of a retail company. If the primary products of a retailer are low gross profits goods, its performance should be lower than other companies with high gross profits. Therefore, we usually evaluate a retailer's performance by using some market measures and the company's product mix structure.

- Comparison Analysis with Competitors

We usually do the comparison analysis within the same retail format. For example, if a retailer's format is a chain store, we usually compare all the chain store companies in terms of the measures I mentioned before. We think it's very important, since this comparison analysis will help us to evaluate a company's market price more correctly.

Interview (Investor – 2)

Interviewee: Retail Industry Analyst in a Taiwanese Securities Exchange Company

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 28 July, 2004 (1000 ~ 1030)

Q1. What are the most important factors relative to your investment decision to a retail company? Why do you think they are important??

We evaluate a retailer's performance with two different analyses: financial analysis and operational analysis. For the financial analysis, we focus on a retail company's expense control, cash flow operation and market measures. For the operational analysis, we concentrate on market share, logistic strength and inventory management and international expansion.

Q2. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

- **<u>Financial Analysis</u>**

- **Expense Control**

Since retail is characterized by low margins, expense control is very important. Therefore, we will examine a retail company's expense control situation before we make any investment decisions.

- **Cash Flow Operation**

We prefer to invest in a retail company that uses its cash flow in fixed assets. Why? Because this company pays more attention to its primary business line. If a retail company invests its cash flow in other companies that do not have any relationship with the retail company, we will regard this retail company as a risky company. The cash flow operation of a company will affect our investment decision.

- **Market Measure**

We usually use P/E ratio be out main measure in order to evaluate a company's market value.

- **<u>Operational Analysis</u>**

- **Market Share**

I think scale is also an important factor for evaluating a retail company's performance. We evaluate a retail company's market share by sales. Furthermore, we will also look at a retail company's growth in store numbers, since we think the amount of stores is key for increasing the sales of a retail company.

- **Logistic Strength and Inventory Management**

Logistics system plays an important role in the retail industry. Therefore, the completeness of the logistics system is also a significant factor for us in evaluating a retail company's performance. How can we examine a retail company's logistic strength? We can evaluate it by checking a retail company's inventory management. We usually use inventory turnover be our main measure to this issue.

- **International Expansion**

A good retail company usually has future international expansion plans, since this company knows that the demand of the domestic market will be saturated one day. Therefore, we will also evaluate a retail company's international expansion plan in order to understand a retail company's future plan.

Interview (Investor – 3)

Interviewee: Ms. Retail Industry Analyst in a Taiwanese Investment Institutite

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 28 July, 2004 (1030 ~ 1100)

Q1. What are the most important factors relative to your investment decision to a retail company? Why do you think they are important??

Before we make an investment decision to a retail company, we usually consider six factors: market share, customer loyalty, cash flow operation, management ability and vision, technology and inventory management.

Q2. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

- **<u>Market Share</u>**

We evaluate a retail company's market share by sales and store numbers. A retail company with large store numbers has an advantage in that all the stores can share fixed costs. Sales growth is also very important in the retail industry, since most retail products have low gross margin. Therefore, before we make an investment decision, we usually look at a retail company's sales and store numbers under a comparative base with other retail companies.

- **<u>Customer Loyalty</u>**

A retail company is usually in direct contact with its customers. Customer loyalty is more important in the retail industry than other industries. We usually use market share to measure a retail company's customer loyalty strength.

- **<u>Cash Flow Operation</u>**

We prefer to invest in a retail company using cash flow for its main business line.

Because this is good for its future development. If a company invests its cash flow to short-term investment in the stock market, the company is more risky. Thus, we usually examine a retail company's cash flow operation before we make an investment decision.

- **Management Ability and Vision**

Every company needs to have the right operational direction. This would depend on the vision of the leader in the company. A leader should have ability to lead its company toward to the right direction. Thus, management ability and vision are also very important performance factors in the retail industry. How can we measure them? We usually measure them by interviewing with management in the company. From the interview, we can obtain information of this company's future plan (vision) and past performance (ability).

- **Technology**

The target customers of a retail company are the general marketplace, and customer demand changes very quickly. How to control customer demand is a very important issue in the retail industry. A retail company can get information about customer demand via its technology system. For example, a retail company can get information about customer structure via a POS system. Therefore, we will also examine a retail company's technology system for investment considerations.

- **Inventory Management**

How to manage a retail company's inventory in order to avoid the dead stock is very important in the retail industry, especially for the food retailers. It depends on a good inventory management system. We usually use inventory turnover to measure the strength of inventory management in a retail company.

Interview (Investor – 4)

Interviewee: Retail Industry Analyst of Taiwan National Investment Trust

Interviewer: Mr. Yu-Chiang Allen Hu (PhD student in University of Edinburgh)

Time: 28 July, 2004 (0900 ~ 0930)

Q1. What are the most important factors relative to your investment decision to a retail company? Why do you think they are important??

We usually evaluate a retail company's performance via three different aspects: financial aspect, operational aspect and macro economical aspect.

Q2. How do you measure the impacts of these factors? What kind of measures you will choose in order to assess these impacts?

- **Financial Aspect**

- **Income Statement Analysis**

With regards to the income statement analysis, we usually do look at four measures: sales and profits growth, gross margin, net income growth, operating expense rate.

- **Sales and Profits Growth**

These two measures are related to a company's market share. We also evaluate a retail company's sales per store or profits per store in order to understand each company's store strength.

- **Gross Margin**

A retail company cannot maintain its gross margin means that the profitability of the products in this company is not high. Under these circumstances, the retail company has to develop new products to increase profits. Therefore, we will also evaluate a retail company's ability to maintain its gross margin.

- **Operating Expense Rate**

There are two main methods to increase a company's net income. One is to increase gross profits; another is to decrease operating expenses. We will also try to understand how a retail company control its operating expenses.

- **Cash Flow Analysis**

If a retail company's cash flow is outflow, we will regard this company as a risky company. Because they do not have any cash flow on hand. However, we will also try to understand the purpose of the cash flow usage. If a retail company use its cash flow to its main business, it will have long-term benefit. Under this situation, cash outflow may not have negative impact.

- **Market Measure Analysis**

We usually use P/E ratio as our market measure in order to evaluate a retail company's market value. What is the rational P/E ratio level? By referring foreign retail company's P/E ratio, we can decide the rational P/E ratio level of a retail company in Taiwan. We will also consider P/E ratio before we make an investment decision.

- **Operational Aspect**

With regards to the operational aspect, we will focus on the management ability, vision and ambition. How can we know the ability, vision and ambition of a company's management? By interviewing them. We usually concentrate on a retail company's future plan, since it will affect this company's future profitability and stability. Moreover, through interviews, we can also understand the personality of a retail company's leader. Such information will also have great impact on our investment decision.

- **Macro Economical Aspect**

Actually, the macro economical impact is not very important in the retail industry, since it is a kind of systematic risk. However, we will also examine some important macro economical measures in order to predict the trend of retail industry development. For example, if the house building industry is in a boom, the sales of the DIY furniture company will also increase. Therefore, macro economical factors are also important performance factors in the retail industry.

Appendix B: Pilot Interview Transcriptions and Reflections

Pilot Interview Transcription	
● Date: 22 December, 2003 ● Time Consuming: about 25 minutes. Participant: Manager of International Department, a Leading Chain Store in Taiwan	
Interviewer	The purpose of this interview is to get information about "How can a retail company performance be measured?" As we know, there are many factors we can use to measure a retail company's performance. But I am not sure what kinds of factors are important. Moreover, if they are important, how can we measure them? The interview is absolutely confidential. Please answer questions based on your experience, there are no right or wrong answers during this discussion. If you feel uncomfortable or you don't want to answer any question, you can choose not to answer the question at any time. Do you have any question?
Participant	No
Interviewer	Thank you. Let's start the interview now.
Participant	OK
Interviewer	How important do you think the factor of "market share" is for measuring performance of retail companies?
Participant	Market share?
Interviewer	Yes
Participant	Measure a company's performance?
Interviewer	Yes
Participant	Market share for measuring performance is important because retailing is about getting consumer's dollar. Because you got a lot of competition, based on the market share you have. You can know how successful you are.
Interviewer	Do you mean market share is a very important factor for measuring a retail company's performance?
Participant	I think it will reflect how well it is performing.
Interviewer	OK, I see
Participant	It's a kind of consequence. It is a result more than a factor for measuring performance.
Interviewer	If you want assess market share, what should we consider?
Participant	What should we consider?
Interviewer	Yes, about market share
Participant	There are reports on the magazines, and usually by sales, and also depending on the store number we have.
Interviewer	But there are something we have to think about market share… For example, if two companies have the same market share, but maybe one company is more risky than the other. Do you think there is something we have to consider?
Participant	Market share? I think market share is just a consequence… For example, why market share is important to my company, because it's a chain store operation. You know we need a certain size in order to... for example to have a centralize distribution system, all of that we need to have a certain market size.

Interviewer	OK, thank you for your answer. And the second question is: How do you feel about the factor of "location"?
Participant	Location for??
Interviewer	For the store…
Participant	Yes. What should I think of location in terms of what?
Interviewer	Yes, I think location is also a very important factor for measuring a retail company's performance...
Participant	Oh... for the performance... The location in terms of performance… I don't think it's a factor for performance assess… Because it just depends on your positioning. For example, our positioning is the convenient service… so we need to close to the customers. If I work in the hypermarket business, we don't need to be on every corner. So, it's not really for performance…
Interviewer	So… do you mean different retailer format will have different location?
Participant	Yes… Every one in the chain store business, you want to choose a better location than your competitors… It means that you like to choose a place with traffic… Because there will have impact of sales.
Interviewer	I understand your opinion, but I want to know how you can decide the location is important…
Participant	The amount of traffic, many people walk around… For example, it could be close to school, underground…
Interviewer	So, do you mean you will consider the population density?
Participant	Yes, and also the rent… If the rent is too high, it will increase the cost…
Interviewer	So you will consider the cost…
Participant	Yes, our selection sometimes also think of… you know… sometimes the stores are very close to each other… so, sometimes we think of opening two stores with the same brand, and close to each other… Because the added sales is higher than a single store…so maybe the previous store, because it is only one store, it's sales will drop…
Interviewer	Yes, that's right...
Participant	But the accumulate sales will increase…
Interviewer	OK, I see… thank you. The third question is there are many factors will affect the stability of future cash flow, such as fashion, and seasonality. How do you measure their influences?
Participant	Of what? Which one?
Interviewer	Such as, seasonality. Because it will cause the volatility of future cash flow...
Participant	We have merchandising people, they deal with that… and every quarter we have performance review… So it usually allows us to change anything in order to increase the sales. If we have too much inventories, we can have a promotion…
Interviewer	Thank you; the next question is what is your opinion of the factor of "sales growth"?
Participant	Sales growth?
Interviewer	Yes, because I think there is a very important feature of retail industry is "low margin"… so if you want to increase your profit, I think the sales is the key point, right?
Participant	Do you mean how to increase sales? There are two methods we can use to increase the sales. One is to increase the margin; the other is to increase the amount of sales. Usually, we play with these two factors; we try to find the balance. There are a lot of strategies in terms of that, for example, if we intend to have a promotion, the margin will be lower. So we have to predict the number of products we have to sell in order to get the margin back.

Interviewer	OK, thank you. The next question is what do you think about the factor of "brand strength"?
Participant	Do you mean like 7-eleven brand? Logo or product brand?
Interviewer	Such as Tesco or Sainsbury…
Participant	Oh, I see… there is value of brand… since customer will have loyalty. People can trust what they sell…
Interviewer	You just mentioned that the customer loyalty is very important. How can you measure the customer loyalty?
Participant	We measure how frequent they go to our store. But I don't have the figure; you can ask the marketing department.
Interviewer	OK. Sorry, but I think if you just measure the frequency of customers visit your store… It's not objective, since they may not pay their money…
Participant	But I think customers usually don't go to convenient store just for looking around, it does not like customers to go to department stores… they just want to look around and don't want to buy anything…
Interviewer	OK, I see… Can you discuss about the importance of Logistics function in retail industry?
Participant	Yes, I think it's very important in the business, since we need our products can be sent to the store at a certain time. For example, for fresh food, we need to order twice a day. So, it's quite frequent. If we don't have a strong delivery system, we may have a shortage…
Interviewer	Do you think "Technology" is a very important factor for measuring performance of a retail company?
Participant	Yes, it's important. In the store level, because the store managers they don't have enough time to do ordering job. If we can give them an IT tool, it can help them to save time. We can also get information from the technology system, such as sales information. They can use the information to forecast how much they can sell for next week.
Interviewer	OK, I see. How can you compare the value of technology between two companies?
Participant	I think the more on time the information you get, the better the decision you will make. I know some of the countries; their data are not immediate. But in Taiwan, we can get our information every day… It's quite quick.
Interviewer	Thank you, The last question… What is your opinion about the factor of "innovation"? Do you think innovation is a very important factor for measuring a retail company's performance?
Participant	Yes… I think innovation is a very important factor, because the retail market is always changing. And we need to continue innovate in order to face the customer's need.
Interviewer	Do you think is it an advantage to a retail company, which always plays the first mover role?
Participant	Yes, I just want to mention that. Because if you always try to innovate, it means that you always try to get larger market share. So, it is an advantage if you always try to keep a head. However, if the market is very close to mature, it's very difficult to make difference.
Interviewer	I see. I agree with you. It's very important for a retail company try to close the customer's demand.
Participant	Yes
Interviewer	All right. Thank you very much for your time and corporation. Your information is very useful for this research. We will analyze the research and give you the feedback in the future. Thank you again!

Pilot Interview Reflections

1. Introduction

This interview is a pilot individual interview for my research on performance rating of retail industry. In studying this topic, both financial and non-financial factors come into play. The purpose of this interview is to get information about *non-financial* factors and how they can be quantified. The participant is previously a junior manager in the International Department of a retail company in Taiwan. The interview is made face-to-face with open questions, and recorded.

2. Analysis of the Interview Process

Reflecting on the literature regarding how to conduct an interview, I have found some strengths and weaknesses in my interview. In the table below, I have listed those criteria found in the literature that were relevant to my interview. For each of the criteria, I evaluated myself on a 5-point Liker scale.

Interview Evaluation Table

Kvale, S. (1996) Interviews: An Introduction to Qualitative Research Interviews Davis and Cosenza (1993) Business Research for Decision Making					
Criteria	Very bad	Not good	Average	Good	Very good
A. Knowledgeable					V
B. Structuring			V		
C. Clear	V				
D. Gentle / Trustworthy			V		
E. Sensitive			V		
F. Open		V			
G. Steering				V	
H. Critical			V		
I. Remembering	V				
J. Interpreting			V		
Patton, M. Q. (1990) Qualitative Evaluation and Research Methods Cooper and Emory (1995) Business Research Methods					
Criteria	Very bad	Not good	Average	Good	Very good
K. Asking Open-ended Questions			V		
L. Avoiding Dichotomous Questions			V		
M. Using Presupposition Questions	V				
N. Asking Singular Questions				V	
O. Using Illustrative Examples		V			
P. Using Probes			V		
Q. Using Announcements	V				
R. Providing Reinforcement				V	
S. Neutrality			V		
T. Tape-recording Issues			V		

Following are some reflections on this self-evaluation.

A. Knowledgeable

I think I showed good knowledge of the topics in the interview due to my previous working experience in the retail sector and my major in finance.

B. Structuring

I introduced the interview very well, remembering to explain clearly the purpose of the interview and tell the Participant that the conversation will be confidential. Furthermore, I rounded off the interview by thanking the participant and promising to provide feedback in the future. This was good.

However, throughout the interview, I did not link the questions with each other to make the conversation flow naturally. For example, after I got a response, I just said, "*OK, thank you. The next question is…*" The main reason was that I was not sure I understood the answer, and instead of expanding when I needed to, I jumped to the next question. Moreover, in designing the questions, I did not consider sequencing the questions. For instance, I could have started with easier or more direct questions and then move on to more analytical ones.

C. Clear

I did not ask very clear questions. This could be seen from the participant frowning and the fact that she continuously tried to confirm my questions.

Example 1:

Interviewer	*How important do you think the factor of "market share" is for measuring performance of retail companies?*
Participant	*Market share?*
Interviewer	*Yes*
Participant	*Measure a company's performance?*
Interviewer	*Yes*

Example 2:

Interviewer	*OK, thank you for your answer… And the second question is... How do you feel about the factor of "location"?*
Participant	*Location for??*
Interviewer	*For the store…*
Participant	*Yes... What should I think of location in terms of what?*

I could have clarified the questions more, but I did not. Furthermore, in example 2, my question was too short and incomplete. I assumed the participant knew I was asking about performance measurement. Another mistake I made was to use professional jargon such as "*the volatility of future cash flow*". I think since the participant is not a finance specialist, she did not know what to answer.

D. Gentle / Trustworthy

I think I was relaxed, polite and interested. My tone of voice was not too loud and I used some gestures to stress certain points. I also tried to maintain eye contact so that the participant may feel respected. Overall, I did not make the participant feel intimidated.

However, because I was eager to get a certain response, I may have seemed provocative in the way I expressed myself sometimes. For example, I challenged an answer by saying, "*OK. Sorry, but I think if you just measure the frequency of customers visit your store... It's not objective, since they may not pay their money...*", as if the Participant did not know what she was saying. I feel this is disrespectful.

E. Sensitive

I did not pay much attention to this aspect when interviewing, but on the whole, I tried to listen to what the participant was saying, and expressed that I understood what was being said, even when I did not.

F. Open

I was not open enough to new ideas on market share. In my opinion, market share is a factor for measuring performance, but the participant said, "*it's a kind of consequence. It's a result more than a factor for measuring performance.*" I did not follow up on the participant's opinion. Instead, I tried to pull the participant over to my point of view. The participant had to repeat again what she thought, and still, I did not seek to follow up but jumped to the next question. I think I need to keep in mind the importance of being open for future interviews.

G. Steering

Basically, I was able to control the interview process. When the Participant digressed, I managed to refocus the interview. For example:

Example 3:

Participant	*Yes... Every one in the chain store business, you want to choose a better location than your competitors... It means that you like to choose a place with traffic... Because there will have impact of sales..*
Interviewer	*I understand your opinion, but I want to know how you can decide the location is important...*

H. Critical

In terms of being critical of what is being said, I would confirm what was said by asking the same question in a different way. For example, the question "*How important do you think the factor of market share is for measuring the performance of retail companies?*" was followed up by "*if a company has a huge market share, is it a positive factor of performance or not?*" However, I did not seek validation every time.

I. Remembering

"Remembering" refers to recalling earlier statements and asking the Participant to elaborate them. I did not do this at all in the interview. With "remembering", I could have linked different parts of the interview together and have more complete answers. I could also have used "remembering" to test reliability and validity of the response.

J. Interpreting

Interpreting refers to clarifying what was said by paraphrasing. In the interview, I did not do this very often. An example of interpreting is as follows.

Example 4:

Participant	*The amount of traffic, many people walk around... For example, it could be close to school, underground...*
Interviewer	*So, do you mean you will consider the population density?*

K. Asking Open-ended Questions

The questions were designed to be open-ended since this research is at the exploratory stage. Most questions started with "*What do you think about...?*" or "*What is your opinion of...?*" Unfortunately, there were occasions where I started with an open question, but in trying to clarify it for the participant, I gave away my opinion. Thus, the open question became a leading question, that is, no longer very open. Here is an example:

Example 5:

Interviewer	*Thank you; the next question is what is your opinion of the factor of "sales growth"?*
Participant	*Sales growth?*
Interviewer	*Yes, because I think there is a very important feature of retail industry is "low margin"... so if you want to increase your profit, I think the sales is the key point, right?*

L. Avoiding Dichotomous Questions

In the beginning of the interview, I avoided dichotomous questions, or questions that lead to "*yes*" or "*no*" answers. However, in the second half of the interview, I asked some dichotomous questions. The mistake was to start the questions with "*do you think...*?". For example: "*Do you think Technology is a very important factor for measuring the performance of a retail company*?" Dichotomous questions are best avoided because the point of doing an interview is to get a rich view of the situation. Also, the Participant is uncooperative; maybe his/her answer would only be "*yes*" or "*no*".

M. Using Presupposition Questions

I did not use presupposition questions in this interview. Nevertheless, I now feel that presupposition questions would be useful for elements that are commonly known to be important factors of performance measurement. For example, sales growth is important to every company. It is a very basic goal of a company. I should not need to ask the participant whether or not it is important, but how important it is.

N. Asking Singular Questions

Most of the time, I only asked singular questions, since this increases clarity. There was one time I asked a multiple question and the response of the Participant reflected confusion on what to answer.

Example 6:

Interviewer	*OK, I see... thank you, the third question is there are many factors will affect the stability of future cash flow, such as fashion, and seasonality. How do you measure their influences?*
Participant	*Of what? Which one?*

This shows that it is indeed very important to ask singular questions.

O. Using Illustrative Examples

I attempted to use illustrative examples to clarify my question, but I did not give very good examples. A good illustrative example should include opposing extremes so that the interviewer does not lead the Participant. For example, in this interview, this is the way I expressed my illustrative example:

Example 7:

Interviewer	*For example, if two companies have the same market share, but maybe one company is more risky than the other. Do you think there is something we have to consider?*

A non-leading way of using illustrative examples would be:

Example 8:

Interviewer	*I've heard some people say that market share is important for measuring performance, but others say that market share is not important for measuring performance. What is your opinion on this?*

P. Using Probes

I used some probes during the interview, but not enough, in my opinion. I could have tried to get deeper insights on some of the topics by using more probes. Example of a probing question:

Example 9:

Participant	*Oh, I see... there is value of brand... since customer will have loyalty. People can trust what they sell...*
Interviewer	*You just mentioned that the customer loyalty is very important. How can you measure the customer loyalty?*

Example of a where I could have added a probing question but did not:

Example 10:

Interviewer	*OK, I see... Can you discuss about the importance of Logistics function in retail industry?*
Participant	*Yes, I think it's very important in the business, since we need our products can be sent to the store at a certain time. For example,[...]*
Interviewer	*Do you think "Technology" is a very important factor for measuring performance of a retail company?*

Here, I just jumped to the next question. I should have probed a bit further on how to *measure* the logistics factor for performance, by asking for instance, "*So, you are saying logistics is important. Could you elaborate on the performance measurements you use in logistics?*"

Q. Using Announcements

I did not introduce the questions in this interview, but just said, "*my next question is...*" I should have used an announcement, such as, "*We've just talked about market share and its relationship to performance. I know you are a chain store operation, so I would also like to ask you about the relationship between performance and store location, if there is one.*"

R. Providing Reinforcement

Overall, I provided encouragements or showed my support to the participant after each question by saying "*thank you*" or "*I see*." At the end of the interview, I stressed how useful the responses were. However, I can improve the way I provide reinforcements.

S. Neutrality

The interviewer should be neutral since there is no right or wrong answer in an interview. The interviewer should show understanding but not express personal opinions. In this interview, I was not always neutral. At times, I gave my own opinion on my own questions, and once, I even said that I agreed with the participant.

T. Tape-recording Issues

There are two areas that I can improve for my next interview. First, before I start recording, I have to give a brief introduction on why I am recording and to reassure the participant on confidentiality of the recording. For example, I could have said, "*I'd like to tape record what you have to say so that I don't miss any of it. It is confidential and only I will listen to it. You can ask me to stop recording at any time.*" Second, the quality of tape recording was not very good and this caused problems for transcribing. Next time, I should place the recording pen nearer to the participant in order to clearly record the answer. Nevertheless, the recording environment chosen was quiet and free from interruptions (The telephone was unplugged). It was suitable for recording.

3. Conclusion

On the whole, my main strength in this interview was my knowledge in the subject matter. I was also good in encouraging the participant to speak. Furthermore, I was in control of the interview process and did not exceed the time limit agreed upon. Also, I mainly asked singular questions. However, even though I asked singular questions, they were not clear enough. I think clarity is the most serious problem I have. If the questions are not clear, the responses will not be clear, and I cannot obtain the information that I really need. I also wasted a lot of time explaining my questions.

Moreover, I scored low for recalling what was said earlier, using announcements and using presupposition questions. All of these elements could have made my questions more interconnected and allowed the whole interview to flow better. The results would have been more comprehensive if I had paid attention to these elements. In terms of being open and using illustrative examples, although I tried to apply these concepts, but I think I was not good enough. I need to be more open to new ideas and follow up on them. I also need to be careful not to lead the participant when using illustrative examples.

Overall, before and during the interview, I should consider each criterion carefully. I need to spend more time on designing the questions, and practice more on how to conduct the interview process. Finally, apart from the criteria in the table above, I discovered from this pilot interview that my questions need some reflection from the part of the participant. They are not issues that most people would think about everyday. For my topic, it would be wise to provide some general questions in advance to the participant so that the participant has time to prepare the answers.

Appendix C: Performance Measures Arrangement

Financial Resources	
Principal	Measures
Profitability	1. EBIT margin 2. EBITDA margin 3. EBITDAR margin 4. Pre-tax profit margin 5. Pre-tax profit on capital 6. Net profit margin 7. Gross profit margin 8. SG&A as % of net sales 9. EBIT on capital 10. Return on total assets 11. Return on total equity 12. Operating margin 13. Dividend payout ratio
Liquidity	14. Current ratio 15. Acid ratio 16. Cash ratio
Sustainability	17. Net operating cash flow / gross Capex 18. Net operating cash flow / maintenance Capex 19. Cash dividend cover 20. Fixed charge cover 21. Interest cover 22. Funds from operations / total debt 23. EBITDA / interest 24. Total debt / discretionary cash flow
Leverage	25. Debt ratio 26. Debt / EBITDA 27. Leased-adjusted net debt / EBITDAR 28. Net debt / market capitalization 29. Total debt / (total debt + market capitalization) 30. Debt to equity ratio
Activity	31. Receivable turnover 32. Inventory turnover 33. Total assets turnover 34. Fixed assets turnover 35. Operating asset turnover
Market Measure	36. P/E Ratio
Financial Scale	37. Net sales 38. Total assets 39. Market share by retail sector (based on sales) 40. Market share by retail sector (based on gross margin) 41. Total capital employed 42. Operation cash flow

Physical Resources	
Principal	Measures
Reach Ability	43. Store numbers 44. Distribution of sales by format and channel 45. The footfalls of major outlets 46. The convenience of transportation and parking 47. Trading area and store locations 48. Size of catchment area in population terms

Legal Resources	
Principal	Measures
Brand Strength	49. Advertising expenses as percentage of sales 50. Trends in sales conversion rate 51. Market capitalization / net assets 52. The sales of private brand products 53. The image of product quality 54. The image of social responsibility 55. The frequency of store layout changing 56. The frequency of marketing strategy redirecting

Human Resources	
Principal	Measures
Human Resource Management	57. Number of payrolls 58. Turnover 59. Communication 60. Performance feedback 61. Absenteeism 62. Staff grievances 63. Staff orientation and training 64. Training effectiveness 65. Internal customer satisfaction, based on quality, timeliness and responsiveness 66. Job satisfaction 67. Employee profit and ownership sharing plan

Organizational Resources	
Principal	Measures
Productivity	68. Saving in energy and communication overheads policy 69. Average weekly sales per square meter 70. Sales conversion rate 71. Spend-per-visit rate 72. Net cash cycle 73. Sales per employee 74. EBIT per employee 75. Sales per human resource cost

Organizational Resources	
Principal	Measures
General Management	76. Internal regulations 77. The annual objectives achievement rate
Expansion Ability	78. The completeness of the franchise system 79. The quality of future expansion plan 80. Store opening program
Organizational Management	81. Empowerment 82. The listening ability of management
Technology Management	83. The investment of technology 84. The strength of data collection and process system
Financial Management	85. Cash flow operation strategy 86. Cost control ability 87. Part-time staffs ratio
Growth Power	88. Sales growth 89. Market value growth 90. Capital growth 91. EBIT growth 92. Number of stores growth 93. Customer footfalls growth 94. The operating income growth 95. Number of payrolls growth
Inventory Management	96. Known loss control 97. Unknown loss control 98. Timeliness and accuracy 99. Efficient warehousing and distribution 100. Out of stock situation
Marketing Management	101. Market positioning 102. The frequency of remodelling 103. Merchandise assortments 104. Pricing policy 105. Differentiate strategy 106. Promotion activities 107. The quality of advertising 108. The number of national brand products 109. The number of local brand products
Product Innovation Ability	110. The amount of new products introduced in a time period 111. The life of new products 112. The speed of new products development 113. The speed of elimination of dead items
Loan Repay Ability	114. Debtor's past credit history 115. Mortgage targets, repay insurance and repay promise 116. Stockholder's background
Diversification	117. Capital expenditures in internet channel 118. Maintaining target customer group in market diversification 119. The consideration of lack of strategic sense 120. The consideration of lack of effective execution 121. Market diversification brings synergy to main business

Informational Resources	
Principal	Measures
Market Segment Risk Management	122. Main market sales as percentage total sales 123. Monthly or quarterly distribution of sales and profits 124. Peak net debt / Average net debt 125. Following fashion trends 126. Facing seasonal demands 127. Awareness of long-term cyclical trends
Strategic Vision	128. Ability to adapt environment change 129. Openness to criticism 130. Willingness to innovate or experiment 131. Recognition of competitive position

Relational Resources	
Principal	Measures
Customer Relations Management	132. Customer complaints management 133. Loyalty card strategy 134. Customer satisfaction 135. Goods returned management
Supplier Relations Management	136. Good global reach 137. Cost sharing with suppliers on promotions 138. Payables turnover
Competitors Relations Management	139. Co-operative alliances opportunity 140. The retailer association

Actions from Customers, Suppliers and Competitors	
Principal	Measures
Customers' Action	141. Changes in customer's preferences or tastes
Suppliers' Action	142. Changes in supplier's relationship, such as changes in contract content
Competitors' Action	143. The innovation from competitors 144. The imitation from competitors

The Societal Resources, the Societal Institutions and the Actions of Government	
Principal	Measures
Political Environmental Factors	145. Government laws and regulations 146. Stability of government 147. The completeness of the infrastructure system 148. The correlation coefficient between government debt / GDP and total sales 149. The correlation coefficient between government avenue / GDP and total sales 150. The correlation coefficient between government expense / GDP and total sales

The Societal Resources, the Societal Institutions and the Actions of Government	
Principal	Measures
Economic Environmental Factors	151. Situations of the global and local market economy 152. Regional economies 153. Situation of the retailing industry 154. The correlation coefficient between GDP and total sales 155. The correlation coefficient between average interest rate and total sales 156. The correlation coefficient between unemployment rate and total sales 157. The correlation coefficient between disposable income and total sales
Socio-culture Environmental Factors	158. Demographic factors (such as, age, sex, material status, household size, education, social class and geographic location) 159. Life style and attitude changes 160. Population structure changes 161. Culture changes (X generation) 162. The correlation coefficient between birth rate and total sales 163. The correlation coefficient between death rate and total sales 164. The correlation coefficient between age structure ratio (0-14years old) and total sales 165. The correlation coefficient between age structure ratio (65 years and above) and total sales
Technological Environmental Factors	166. New products 167. New production process 168. Innovation of new technology equipment 169. New development in information handling 170. The correlation coefficient between total government spending for R&D and total sales

Appendix D: <u>Performance Measures Regrouping (Based on data Availability)</u>

Quantifiable Measure and Available Data Group	
Financial Resources	
Principal	Measures
Profitability	1. EBIT margin 2. EBITDA margin 3. EBITDAR margin 4. Pre-tax profit margin 5. Pre-tax profit on capital 6. Net profit margin 7. Gross profit margin 8. SG&A as % of net sales 9. EBIT on capital 10. Return on total assets 11. Return on total equity 12. Operating margin 13. Dividend payout ratio
Liquidity	14. Current ratio 15. Acid ratio 16. Cash ratio
Sustainability	17. Net operating cash flow / gross Capex 18. Cash dividend cover 19. Fixed charge cover 20. Interest cover 21. Funds from operations / total debt 22. EBITDA / interest 23. Total debt / discretionary cash flow
Leverage	24. Debt ratio 25. Debt / EBITDA 26. Leased-adjusted net debt / EBITDAR 27. Net debt / market capitalization 28. Total debt / (total debt + market capitalization) 29. Debt to equity ratio
Activity	30. Receivable turnover 31. Inventory turnover 32. Total assets turnover 33. Fixed assets turnover
Market Measure	34. P/E Ratio
Financial Scale	35. Net sales 36. Total assets 37. Market share by retail sector (based on sales) 38. Market share by retail sector (based on gross margin) 39. Total capital employed 40. Operation cash flow

Quantifiable Measure and Available Data Group (Con.)	
Physical Resources	
Principal	Measures
Reach Ability	41. Store numbers
Legal Resources	
Principal	Measures
Brand Strength	42. Market capitalization / net assets
Human Resources	
Principal	Measures
Human Resource Management	43. Number of payrolls
Organizational Resources	
Principal	Measures
Productivity	44. Net cash cycle 45. Sales per employee 46. EBIT per employee
Growth Power	47. Sales growth 48. Market value growth 49. Capital growth 50. EBIT growth 51. Number of stores growth 52. The operating income growth 53. Number of payrolls growth
Informational Resources	
Principal	Measures
Market Segment Risk Management	54. Main market sales as percentage total sales
Relational Resources	
Principal	Measures
Supplier Relations Management	55. Payables turnover
The Societal Resources, the Societal Institutions and the Actions of Government	
Principal	Measures
Political Environmental Factors	56. The correlation coefficient between government debt / GDP and total sales 57. The correlation coefficient between government avenue / GDP and total sales 58. The correlation coefficient between government expense / GDP and total sales
Economic Environmental Factors	59. The correlation coefficient between GDP and total sales 60. The correlation coefficient between average interest rate and total sales 61. The correlation coefficient between unemployment rate and total sales 62. The correlation coefficient between disposable income and total sales

Quantifiable Measure and Available Data Group (Con.)	
The Societal Resources, the Societal Institutions and the Actions of Government	
Principal	Measures
Socio-culture Environmental Factors	63. The correlation coefficient between birth rate and total sales 64. The correlation coefficient between death rate and total sales 65. The correlation coefficient between age structure ratio (0-14years old) and total sales 66. The correlation coefficient between age structure ratio (65 years and above) and total sales
Technological Environmental Factors	67. The correlation coefficient between total government spending for R&D and total sales

Quantifiable Measure But No Available Data Group	
Financial Resources	
Principal	Measures
Sustainability	68. Net operating cash flow / maintenance Capex
Activity	69. Operating asset turnover
Physical Resources	
Principal	Measures
Reach Ability	70. Distribution of sales by format and channel 71. The footfalls of major outlets 72. Size of catchment area in population terms
Legal Resources	
Principal	Measures
Brand Strength	73. Advertising expenses as percentage of sales 74. The sales of private brand products 75. The frequency of store layout changing 76. The frequency of marketing strategy redirecting
Human Resources	
Principal	Measures
Human Resource Management	77. Turnover
Organizational Resources	
Principal	Measures
Productivity	78. Saving in energy and communication overheads policy 79. Average weekly sales per square meter 80. Sales conversion rate 81. Spend-per-visit rate 82. Sales per human resource cost
General Management	83. The annual objectives achievement rate
Financial Management	84. Part-time staffs ratio
Growth Power	85. Customer footfalls growth

Quantifiable Measure But No Available Data Group (Con.)	
Organizational Resources	
Principal	Measures
Marketing Management	86. The frequency of remodelling 87. The number of national brand products 88. The number of local brand products
Product Innovation Ability	89. The amount of new products introduced in a time period 90. The life of new products 91. The speed of new products development 92. The speed of elimination of dead items
Diversification	93. Capital expenditures in internet channel
Informational Resources	
Principal	Measures
Market Segment Risk Management	94. Monthly or quarterly distribution of sales and profits 95. Peak net debt / Average net debt

Difficult To Quantify Group	
Physical Resources	
Principal	Measures
Reach Ability	96. The convenience of transportation and parking 97. Trading area and store locations
Legal Resources	
Principal	Measures
Brand Strength	98. Trends in sales conversion rate 99. The image of product quality 100. The image of social responsibility
Human Resources	
Principal	Measures
Human Resource Management	101. Communication 102. Performance feedback 103. Absenteeism 104. Staff grievances 105. Staff orientation and training 106. Training effectiveness 107. Internal customer satisfaction, based on quality, timeliness and responsiveness 108. Job satisfaction 109. Employee profit and ownership sharing plan
Organizational Resources	
Principal	Measures
General Management	110. Internal regulations
Expansion Ability	111. The completeness of the franchise system 112. The quality of future expansion plan 113. Store opening program

Difficult To Quantify Group (Con.)	
Organizational Resources	
Principal	Measures
Organizational Management	114. Empowerment 115. The listening ability of management
Technology Management	116. The investment of technology 117. The strength of data collection and process system
Financial Management	118. Cash flow operation strategy 119. Cost control ability
Inventory Management	120. Known loss control 121. Unknown loss control 122. Timeliness and accuracy 123. Efficient warehousing and distribution 124. Out of stock situation
Marketing Management	125. Market positioning 126. Merchandise assortments 127. Pricing policy 128. Differentiate strategy 129. Promotion activities 130. The quality of advertising
Loan Repay Ability	131. Debtor's past credit history 132. Mortgage targets, repay insurance and repay promise 133. Stockholder's background
Diversification	134. Maintaining target customer group in market diversification 135. The consideration of lack of strategic sense 136. The consideration of lack of effective execution 137. Market diversification brings synergy to main business
Informational Resources	
Principal	Measures
Market Segment Risk Management	138. Following fashion trends 139. Facing seasonal demands 140. Awareness of long-term cyclical trends
Strategic Vision	141. Ability to adapt environment change 142. Openness to criticism 143. Willingness to innovate or experiment 144. Recognition of competitive position
Relational Resources	
Principal	Measures
Customer Relations Management	145. Customer complaints management 146. Loyalty card strategy 147. Customer satisfaction 148. Goods returned management
Supplier Relations Management	149. Good global reach 150. Cost sharing with suppliers on promotions
Competitors Relations Management	151. Co-operative alliances opportunity 152. The retailer association

Difficult To Quantify Group (Con.)	
<u>Actions from Customers, Suppliers and Competitors</u>	
Principal	Measures
Customers' Action	153. Changes in customer's preferences or tastes
Suppliers' Action	154. Changes in supplier's relationship, such as changes in contract content
Competitors' Action	155. The innovation from competitors 156. The imitation from competitors
<u>The Societal Resources, the Societal Institutions and the Actions of Government</u>	
Principal	Measures
Political Environmental Factors	157. Government laws and regulations 158. Stability of government 159. The completeness of the infrastructure system
<u>The Societal Resources, the Societal Institutions and the Actions of Government</u>	
Principal	Measures
Economic Environmental Factors	160. Situations of the global and local market economy 161. Regional economies 162. Situation of the retailing industry
Socio-culture Environmental Factors	163. Demographic factors (such as, age, sex, material status, household size, education, social class and geographic location) 164. Life style and attitude changes 165. Population structure changes 166. Culture changes (X generation)
Technological Environmental Factors	167. New products 168. New production process 169. Innovation of new technology equipment 170. New development in information handling

Appendix E: E-Questionnaire (English Version)

RETAIL INDUSTRY PERFORMANCE MEASUREMENT

Thank you very much for your time and kind participation. The purpose of this questionnaire is to obtain information on factors for measuring the performance of retail companies. It is part of a research project carried out at the University of Edinburgh by Mr. Yu-Chiang Allen Hu (Y.A.HU@sms.ed.ac.uk) as part of a doctoral thesis.

Two sections in this questionnaire:

A) Evaluating the Importance of Various Factors to the Performance of a Retail Company

B) Evaluating Performance Through a Rating Process

Attention:

A) The survey is on a single page. Please use your mouse to scroll down the page.

B) Please answer questions based on your experience, as there are no right or wrong answers in this questionnaire.

C) The questionnaire will take at most 15 minutes to complete with a total of 9 main questions.

D) You need to complete the survey in one session, since the file cannot be saved.

E) The survey is absolutely "Confidential". When you submit, your responses are automatically sent to the researcher only.

Please start by typing the password provided in the email to validate this survey:

Section A: Importance of Various Factors to the Performance of a Retail Company

Section A asks you what you think are the key factors that influences the performance of your company, as well as how important these key factors are.

Question 1:

The following factors are related to Reach Ability, Brand Strength, Human Resource Management, Expansion Ability and Productivity.

How important are these ten factors to your company's performance?

Please tick one box only from 'Don't Know' to 'Absolutely Important' for each statement.

	Absolutely Important	Very Important	Moderately Important	Not Very Important	Don't Know
Number of customer visits	☐	☐	☐	☐	☐
Store location	☐	☐	☐	☐	☐
Sales of the private brand products	☐	☐	☐	☐	☐
Social responsibility	☐	☐	☐	☐	☐
Employee turnover rate	☐	☐	☐	☐	☐
Staff training	☐	☐	☐	☐	☐
Franchise system	☐	☐	☐	☐	☐
Store opening strategy	☐	☐	☐	☐	☐
Sales per store	☐	☐	☐	☐	☐
Spending-per-visit rate	☐	☐	☐	☐	☐

Question 2:

The following factors are related to General Management, Technology Management, Organizational Management, Inventory Management and Marketing Management.

How important are these ten factors to your company's performance?

Please tick one box only from 'Don't Know' to 'Absolutely Important' for each statement.

	Absolutely Important	Very Important	Moderately Important	Not Very Important	Don't Know
Internal procedures	☐	☐	☐	☐	☐

	Absolutely Important	Very Important	Moderately Important	Not Very Important	Don't Know
Achievement of year-end goals	☐	☐	☐	☐	☐
Investments in technology development	☐	☐	☐	☐	☐
Quality of data collection and process system	☐	☐	☐	☐	☐
Empowerment of staff	☐	☐	☐	☐	☐
Response to staff issues	☐	☐	☐	☐	☐
Inventory loss control	☐	☐	☐	☐	☐
Inventory service level	☐	☐	☐	☐	☐
Market positioning	☐	☐	☐	☐	☐
Store renovation/redecoration	☐	☐	☐	☐	☐

Question 3:

The following factors are related to Financial Management, Product Innovation, Loan Repay Ability, Diversification, Market Segment Risk Control and Strategic Vision.

How important are these twelve factors to your company's performance?

Please tick one box only from 'Don't Know' to 'Absolutely Important' for each statement.

	Absolutely Important	Very Important	Moderately Important	Not Very Important	Don't Know
Expense control ability	☐	☐	☐	☐	☐
Percentage of part-time staff	☐	☐	☐	☐	☐
Shelf-life of new products	☐	☐	☐	☐	☐
Speed of new products development	☐	☐	☐	☐	☐
Past credit history	☐	☐	☐	☐	☐
Financial support from stockholders	☐	☐	☐	☐	☐
Internet channel development	☐	☐	☐	☐	☐
Maintaining target customer group in market diversification	☐	☐	☐	☐	☐

	Absolutely Important	Very Important	Moderately Important	Not Very Important	Don't Know
Following fashion trends	☐	☐	☐	☐	☐
Facing seasonal demands	☐	☐	☐	☐	☐
Openness to criticism	☐	☐	☐	☐	☐
Willingness to innovate	☐	☐	☐	☐	☐

Question 4:

The following factors are related to Customer's Relationship, Supplier's Relationship, Competitor's Relationship, and External Environmental Factors.

How important are these twelve factors to your company's performance?

Please tick one box only from 'Don't Know' to 'Absolutely Important' for each statement.

	Absolutely Important	Very Important	Moderately Important	Not Very Important	Don't Know
Customer complaints management	☐	☐	☐	☐	☐
Cost sharing with suppliers for promotions	☐	☐	☐	☐	☐
Joint venture opportunity with competitors	☐	☐	☐	☐	☐
Changes in customer's preferences	☐	☐	☐	☐	☐
Changes in supplier's contract content	☐	☐	☐	☐	☐
The innovation and imitation from competitors	☐	☐	☐	☐	☐
Change of government laws	☐	☐	☐	☐	☐
Innovation of new technology equipment	☐	☐	☐	☐	☐
Change of population structure	☐	☐	☐	☐	☐
Stability of government	☐	☐	☐	☐	☐
New management system software development	☐	☐	☐	☐	☐
Change of lifestyle	☐	☐	☐	☐	☐

Section B: Evaluating Performance Through a Ranking Process

Section B asks you how well do you think your company is doing in the factors mentioned in section A? Or how much impact these factors have on your company's performance?

Question 5:

How well do you think your company is doing in the following factors?

Please tick one box only from 'Don't Know' to 'Extremely good' for each statement.

	Extremely Good	Very Good	Moderately Good	Not Very Good	Don't Know
Number of customer visits	☐	☐	☐	☐	☐
Store location	☐	☐	☐	☐	☐
Sales of the private brand products	☐	☐	☐	☐	☐
Social responsibility	☐	☐	☐	☐	☐
Employee turnover rate	☐	☐	☐	☐	☐
Staff training	☐	☐	☐	☐	☐
Franchise system	☐	☐	☐	☐	☐
Store opening strategy	☐	☐	☐	☐	☐
Sales per store	☐	☐	☐	☐	☐
Spending-per-visit rate	☐	☐	☐	☐	☐
Internal procedures	☐	☐	☐	☐	☐

Question 6:

How well do you think your company is doing in the following factors?

Please tick one box only from 'Don't Know' to 'Extremely good' for each statement.

	Extremely Good	Very Good	Moderately Good	Not Very Good	Don't Know
Achievement of year-end goals	☐	☐	☐	☐	☐

	Extremely Good	Very Good	Moderately Good	Not Very Good	Don't Know
Investments in technology development	☐	☐	☐	☐	☐
Quality of data collection and process system	☐	☐	☐	☐	☐
Empowerment of staff	☐	☐	☐	☐	☐
Response to staff issues	☐	☐	☐	☐	☐
Inventory loss control	☐	☐	☐	☐	☐
Inventory service level	☐	☐	☐	☐	☐
Market positioning	☐	☐	☐	☐	☐
Store renovation/redecoration	☐	☐	☐	☐	☐
Expense control ability	☐	☐	☐	☐	☐
Percentage of part-time staff	☐	☐	☐	☐	☐
Shelf-life of new products	☐	☐	☐	☐	☐

Question 7:

How well do you think your company is doing in the following factors?

Please tick one box only from 'Don't Know' to 'Extremely good' for each statement.

	Extremely Good	Very Good	Moderately Good	Not Very Good	Don't Know
Speed of new products development	☐	☐	☐	☐	☐
Past credit history	☐	☐	☐	☐	☐
Financial support from stockholders	☐	☐	☐	☐	☐
Internet channel development	☐	☐	☐	☐	☐
Maintaining target customer group in market diversification	☐	☐	☐	☐	☐
Following fashion trends	☐	☐	☐	☐	☐
Facing seasonal demands	☐	☐	☐	☐	☐
Openness to criticism	☐	☐	☐	☐	☐
Willingness to innovate	☐	☐	☐	☐	☐

	Extremely Good	Very Good	Moderately Good	Not Very Good	Don't Know
Customer complaints management	☐	☐	☐	☐	☐
Cost sharing with suppliers for promotions	☐	☐	☐	☐	☐
Joint venture opportunity with competitors	☐	☐	☐	☐	☐

Question 8:

How strong impact the following factors have on your company's performance?

Please tick one box only from 'Don't Know' to 'Extremely Strong' for each statement.

	Extremely Strong	Very Strong	Moderately Strong	Not Very Strong	Don't Know
Changes in customer's preferences	☐	☐	☐	☐	☐
Changes in supplier's contract content	☐	☐	☐	☐	☐
The innovation and imitation from competitors	☐	☐	☐	☐	☐
Change of government laws	☐	☐	☐	☐	☐
Innovation of new technology equipment	☐	☐	☐	☐	☐
Change of population structure	☐	☐	☐	☐	☐
Stability of government	☐	☐	☐	☐	☐
New management system software development	☐	☐	☐	☐	☐
Change of lifestyle	☐	☐	☐	☐	☐

Question 9: Additional Comments

If you have any additional comments want to share with us, please type here:

Please choose your job content:

	Job Content
Marketing (including purchasing and advertising)	☐
Operations	☐
Store development (including construction franchising)	☐
Logistics	☐
Accounting and finance	☐
Human resources	☐
Information systems	☐
Other	☐

Thank you very much again for your time and kind assistance.

Your feedback is greatly appreciated.

How to Submit This Questionnaire:

A) Click on "Submit Survey" below. Please click only once. Your e-mail will not be revealed.

B) Wait for a confirmation page to appear.

Submit Survey

Appendix F: E-Questionnaire (Mandarin Version)

零售業績效評估調查

非常感謝您寶貴的時間參與此問卷調查。此問卷調查之目的爲取得有關於零售業績效衡量變數的相關資訊。此問卷隸屬於英國愛丁堡大學管理學院研究專案(University of Edinburgh, Management School)，專案主持人爲財金博士班研究生胡玉強。(Mr. Yu-Chiang Allen Hu; email: Y.A.Hu@sms.ed.ac.uk)

此問卷包含二個主要單元：

A) 衡量不同變數對於零售業績效之重要性。
B) 評估貴公司於每項變數的績效。

注意事項：

A) 此問卷僅有一頁，煩請使用滑鼠移動您的頁面。
B) 煩請基於您本身的經驗或認知來回答問題，此問卷之任何問題並無標準答案。
C) 此問卷無法儲存您已填寫寫過之內容，故請一次填寫完成。
D) 此問卷約需歷時 15 分鐘 填寫 9 個主要問題。
E) 此問卷將嚴厲遵守『保密』原則，當您填寫完成，此問卷將僅會傳送給研究者。

在填寫問卷之前，煩請先填入密碼於下列方格中。(密碼位於寄予貴公司之 Email 中)

第一單元：衡量不同變數對於零售業績效之重要性

第一單元將會詢問您有關一系列績效變數對於貴公司績效的影響程度以及他們對於貴公司績效的重要性。

問題 1：

下列因素攸關於一家零售業公司與顧客接觸的能力、品牌優勢、人力資源管理、擴張能力以及生產力。

您覺得下列因素對於貴公司之績效有多重要？對於每一項績效變數，請選擇從'不確定' 至 '極爲重要'中之一個答案。

	極爲重要	非常重要	重要	稍微重要	不確定
來客數	☐	☐	☐	☐	☐
店面位置	☐	☐	☐	☐	☐
自有品牌產品之銷售額	☐	☐	☐	☐	☐
社會責任	☐	☐	☐	☐	☐
員工離職率	☐	☐	☐	☐	☐
員工教育訓練	☐	☐	☐	☐	☐
加盟系統	☐	☐	☐	☐	☐
開店策略	☐	☐	☐	☐	☐
單店銷售額	☐	☐	☐	☐	☐
顧客每次來店消費額	☐	☐	☐	☐	☐

問題 2：

下列因素攸關於一家零售業公司之一般管理能力、科技管理能力、組織管理能力、存貨管理能力以及行銷管理能力。

您覺得下列因素對於貴公司之績效有多重要？對於每一項績效變數，請選擇從'不確定' 至 '極爲重要'中之一個答案。

	極爲重要	非常重要	重要	稍微重要	不確定
內控制度	☐	☐	☐	☐	☐
年度計畫達成率	☐	☐	☐	☐	☐
科技發展之投資	☐	☐	☐	☐	☐

	極爲重要	非常重要	重要	稍微重要	不確定
資料收集與處理系統之品質	☐	☐	☐	☐	☐
充分授權	☐	☐	☐	☐	☐
管理階層能回應員工需求	☐	☐	☐	☐	☐
存貨盤點制度	☐	☐	☐	☐	☐
商品能及時送達店面	☐	☐	☐	☐	☐
差異化行銷	☐	☐	☐	☐	☐
店面的 Remodeling	☐	☐	☐	☐	☐

問題 3：

下列因素攸關於一家零售業公司之財務管理能力、產品創新、償債能力、多角化策略、市場風險控管以及策略願景。

您覺得下列因素對於貴公司之績效有多重要？對於每一項績效變數，請選擇從‘不確定’ 至 ‘極爲重要’中之一個答案。

	極爲重要	非常重要	重要	稍微重要	不確定
費用控管	☐	☐	☐	☐	☐
兼職員工比率	☐	☐	☐	☐	☐
新產品之上架時間	☐	☐	☐	☐	☐
新產品開發速度	☐	☐	☐	☐	☐
公司過去債信記錄	☐	☐	☐	☐	☐
股東之財務支持	☐	☐	☐	☐	☐
網際網路銷售管道的發展	☐	☐	☐	☐	☐
多角化之後仍能維持原主要顧客群	☐	☐	☐	☐	☐
迎合顧客追求流行之需求	☐	☐	☐	☐	☐
面對季節性的需求	☐	☐	☐	☐	☐
公司能夠接受外界批評	☐	☐	☐	☐	☐
公司之創新意願	☐	☐	☐	☐	☐

問題 4：

下列因素攸關於一家零售業公司與顧客、供應商以及競爭者之關係，以及外在環境績效變數。

您覺得下列因素對於貴公司之績效有多重要？對於每一項績效變數，請選擇從'不確定' 至 '極爲重要'中之一個答案。

	極爲重要	非常重要	重要	稍微重要	不確定
顧客抱怨管理	☐	☐	☐	☐	☐
與供應商分享促銷活動成本	☐	☐	☐	☐	☐
與競爭者策略聯盟的機會	☐	☐	☐	☐	☐
消費者偏好之改變	☐	☐	☐	☐	☐
與供應商契約內容之改變	☐	☐	☐	☐	☐
競爭者之創新與模仿	☐	☐	☐	☐	☐
政府法令的改變	☐	☐	☐	☐	☐
設備的創新，如 POS 系統	☐	☐	☐	☐	☐
人口結構的改變	☐	☐	☐	☐	☐
政府政策的穩定性	☐	☐	☐	☐	☐
新管理軟體系統發展	☐	☐	☐	☐	☐
消費者生活型態的改變	☐	☐	☐	☐	☐

第二單元：評估貴公司於每項變數的績效

第二單元將會詢問您針對第一單元所列示的所有績效變數，您認爲貴公司的表現如何？或者是這些績效變數對於貴公司的營運績效影響有多大？

問題 5：

您覺得貴公司於下列績效變數之表現如何？ 對於每一項績效變數，請選擇從'不確定' 至 '極佳'中之一個答案。

	極佳	非常好	還可以	不是很好	不確定
來客數					
店面位置					
自有品牌產品之銷售額					
社會責任					
員工離職率					
員工教育訓練					
加盟系統					
開店策略					
單店銷售額					
顧客每次來店消費額					
內控制度					

問題 6：

您覺得貴公司於下列績效變數之表現如何？ 對於每一項績效變數，請選擇從'不確定' 至 '極佳'中之一個答案。

	極佳	非常好	還可以	不是很好	不確定
年度計畫達成率					
科技發展之投資					
資料收集與處理系統之品質					
充分授權					
管理階層能回應員工需求					
存貨盤點制度					
商品能及時送達店面					
差異化行銷					
店面的 Remodeling					
費用控管					

	極佳	非常好	還可以	不是很好	不確定
兼職員工比率	☐	☐	☐	☐	☐
新產品之上架時間	☐	☐	☐	☐	☐

問題 7：

您覺得貴公司於下列績效變數之表現如何？ 對於每一項績效變數，請選擇從 '不確定' 至 '極佳'中之一個答案。

	極佳	非常好	還可以	不是很好	不確定
新產品開發速度	☐	☐	☐	☐	☐
公司過去債信記錄	☐	☐	☐	☐	☐
股東之財務支持	☐	☐	☐	☐	☐
網際網路銷售管道的發展	☐	☐	☐	☐	☐
多角化之後仍能維持原主要顧客群	☐	☐	☐	☐	☐
迎合顧客追求流行之需求	☐	☐	☐	☐	☐
面對季節性的需求	☐	☐	☐	☐	☐
公司能夠接受外界批評	☐	☐	☐	☐	☐
公司之創新意願	☐	☐	☐	☐	☐
顧客抱怨管理	☐	☐	☐	☐	☐
與供應商分享促銷活動成本	☐	☐	☐	☐	☐
與競爭者策略聯盟的機會	☐	☐	☐	☐	☐

問題 8：

您覺得下列因素對貴公司績效之影響有多強烈？對於每一項績效變數，請選擇從 '不確定' 至 '極強烈' 中之一個答案。

	極強烈	很強烈	強烈	稍微強烈	不確定
消費者偏好之改變	☐	☐	☐	☐	☐

	極強烈	很強烈	強烈	稍微強烈	不確定
與供應商契約內容之改變	☐	☐	☐	☐	☐
競爭者之創新與模仿	☐	☐	☐	☐	☐
政府法令的改變	☐	☐	☐	☐	☐
設備的創新，如 POS 系統	☐	☐	☐	☐	☐
人口結構的改變	☐	☐	☐	☐	☐
政府政策的穩定性	☐	☐	☐	☐	☐
新管理系統軟體發展	☐	☐	☐	☐	☐
消費者生活型態的改變	☐	☐	☐	☐	☐

問題 9：其他建議

如果您對於本問卷調查有更多的意見或建議，煩請於下列空格中填寫，謝謝。

請選擇您服務的單位

	服務單位
行銷部門 (包含採購與廣告單位)	☐
營業部門	☐
店面發展部門 (包含加盟以及工程單位)	☐
物流配送部門	☐
財務部門	☐
人力資源部門	☐
電腦系統部門	☐
其他	☐

再次感謝您寶貴的時間與協助。

您的回饋將使本研究更加完善且客觀！

如何回覆此問卷？

A) 煩請點選左下角"Submit Survey"。請點選一次即可。
B) 請等待確認頁之出現。

Submit Survey

Powered by SurveySolutions XP: Conduct your own online surveys

Appendix G: <u>Survey Descriptive Analysis (Mean and Median Data)</u>

Importance Rank based on the Department											
Marketing (including Advertising)				Operations and Store Development				Accounting, Finance and Auditing			
Code	Mean	Code	Median	Code	Mean	Code	Median	Code	Mean	Code	Median
V19	1.40	V19	1	V32	1.22	V1	1	V1	1.26	V1	1
V32	1.40	V32	1	V33	1.22	V2	1	V2	1.26	V2	1
V33	1.50	V33	1	V19	1.33	V6	1	V32	1.36	V7	1
V1	1.60	V2	1.5	V9	1.39	V7	1	V19	1.41	V10	1
V2	1.60	V1	2	V18	1.42	V9	1	V33	1.46	V11	1
V44	1.60	V3	2	V2	1.44	V18	1	V10	1.54	V19	1
V29	1.65	V5	2	V7	1.44	V19	1	V11	1.54	V29	1
V30	1.70	V6	2	V29	1.44	V29	1	V44	1.54	V30	1
V18	1.75	V7	2	V30	1.48	V30	1	V6	1.56	V32	1
V28	1.80	V8	2	V6	1.50	V32	1	V30	1.56	V33	1
V36	1.80	V9	2	V1	1.56	V33	1	V7	1.59	V44	1
V10	1.90	V10	2	V24	1.67	V3	2	V18	1.59	V3	2
V23	1.90	V12	2	V31	1.69	V4	2	V28	1.59	V4	2
V24	1.95	V13	2	V36	1.72	V5	2	V14	1.62	V5	2
V6	2.00	V14	2	V10	1.74	V8	2	V8	1.64	V6	2
V7	2.00	V15	2	V11	1.74	V10	2	V9	1.67	V8	2
V8	2.00	V16	2	V40	1.74	V11	2	V29	1.67	V9	2
V9	2.00	V18	2	V8	1.78	V12	2	V21	1.74	V12	2
V14	2.00	V20	2	V44	1.78	V13	2	V17	1.79	V13	2
V31	2.10	V21	2	V21	1.81	V14	2	V39	1.79	V14	2
V41	2.10	V23	2	V15	1.85	V15	2	V12	1.85	V15	2
V21	2.11	V24	2	V14	1.85	V16	2	V16	1.85	V16	2
V12	2.20	V26	2	V16	1.85	V17	2	V41	1.87	V17	2
V16	2.20	V27	2	V12	1.93	V20	2	V36	1.88	V18	2
V26	2.20	V28	2	V28	1.96	V21	2	V24	1.90	V20	2
V39	2.20	V29	2	V41	2.04	V23	2	V31	1.92	V21	2
V15	2.25	V30	2	V38	2.06	V24	2	V40	1.95	V23	2
V42	2.30	V31	2	V39	2.07	V25	2	V38	1.96	V24	2
V3	2.35	V34	2	V4	2.11	V26	2	V23	1.97	V25	2
V20	2.35	V36	2	V34	2.19	V27	2	V34	2.03	V26	2
V27	2.35	V38	2	V17	2.22	V28	2	V15	2.08	V27	2
V40	2.35	V39	2	V5	2.23	V31	2	V25	2.08	V28	2
V5	2.37	V40	2	V37	2.28	V34	2	V4	2.08	V31	2
V11	2.40	V41	2	V20	2.30	V36	2	V42	2.10	V34	2
V13	2.40	V42	2	V27	2.30	V37	2	V43	2.26	V36	2
V34	2.40	V44	2	V3	2.31	V38	2	V5	2.28	V37	2
V25	2.45	V11	2.5	V13	2.38	V39	2	V20	2.28	V38	2
V38	2.46	V25	2.5	V25	2.41	V40	2	V37	2.29	V39	2
V17	2.53	V37	2.5	V43	2.41	V41	2	V26	2.31	V40	2
V37	2.57	V4	3	V26	2.48	V42	2	V3	2.38	V41	2
V35	2.62	V17	3	V42	2.48	V43	2	V27	2.44	V42	2
V43	2.65	V22	3	V23	2.52	V44	2	V13	2.45	V43	2
V4	2.73	V35	3	V35	2.81	V22	3	V35	2.83	V22	3
V22	2.85	V43	3	V22	2.81	V35	3	V22	3.00	V35	3

Importance Rank based on the Department (Con.)											
Human Resources				**Investor Relations**				**R&D and Information System**			
Rank	Mean	Rank	Median	Rank	Mean	Rank	Median	Rank	Mean	Rank	Median
V1	1.27	V1	1	V7	1.00	V1	1	V2	1.32	V2	1
V2	1.45	V2	1	V1	1.52	V2	1	V32	1.53	V32	1
V9	1.60	V19	1	V2	1.52	V7	1	V1	1.63	V1	2
V18	1.73	V32	1	V32	1.58	V8	1	V19	1.63	V4	2
V19	1.73	V9	1.5	V33	1.58	V10	1	V44	1.63	V6	2
V10	1.80	V6	2	V10	1.58	V30	1.5	V30	1.68	V7	2
V32	1.82	V7	2	V12	1.62	V32	1.5	V33	1.68	V8	2
V33	1.82	V8	2	V30	1.65	V33	1.5	V29	1.74	V9	2
V7	1.91	V10	2	V6	1.69	V3	2	V7	1.78	V10	2
V8	1.91	V12	2	V8	1.72	V5	2	V9	1.79	V11	2
V30	1.91	V14	2	V21	1.81	V6	2	V6	1.84	V12	2
V6	2.00	V18	2	V19	1.85	V11	2	V8	1.84	V13	2
V21	2.00	V21	2	V18	1.96	V12	2	V11	1.84	V14	2
V36	2.00	V23	2	V20	1.96	V13	2	V18	1.84	V15	2
V44	2.00	V24	2	V31	2.00	V14	2	V36	1.86	V16	2
V14	2.09	V29	2	V13	2.04	V15	2	V14	1.89	V17	2
V24	2.09	V30	2	V14	2.04	V16	2	V40	1.89	V18	2
V29	2.18	V33	2	V11	2.12	V18	2	V24	1.95	V19	2
V38	2.22	V34	2	V16	2.12	V19	2	V41	1.95	V20	2
V26	2.30	V36	2	V44	2.15	V20	2	V39	2.00	V21	2
V34	2.36	V38	2	V15	2.19	V21	2	V16	2.05	V23	2
V40	2.36	V39	2	V41	2.27	V26	2	V23	2.05	V24	2
V41	2.36	V40	2	V3	2.28	V28	2	V42	2.05	V25	2
V16	2.40	V41	2	V5	2.35	V31	2	V10	2.11	V26	2
V42	2.40	V44	2	V28	2.35	V34	2	V12	2.11	V27	2
V25	2.44	V26	2.5	V26	2.44	V41	2	V28	2.11	V28	2
V12	2.45	V42	2.5	V43	2.52	V43	2	V31	2.11	V29	2
V39	2.45	V3	3	V40	2.54	V44	2	V17	2.14	V30	2
V15	2.50	V4	3	V24	2.58	V24	2.5	V15	2.21	V31	2
V23	2.55	V5	3	V34	2.60	V40	2.5	V34	2.21	V33	2
V35	2.56	V11	3	V39	2.62	V22	3	V27	2.22	V34	2
V17	2.60	V13	3	V25	2.64	V23	3	V21	2.26	V36	2
V31	2.60	V15	3	V27	2.72	V25	3	V25	2.26	V39	2
V37	2.60	V16	3	V29	2.76	V27	3	V4	2.29	V40	2
V3	2.70	V17	3	V23	2.77	V29	3	V13	2.32	V41	2
V28	2.73	V20	3	V42	2.84	V39	3	V26	2.37	V42	2
V27	2.78	V22	3	V22	3.08	V42	3	V43	2.37	V43	2
V11	2.82	V25	3	V4	NA	V4	NA	V20	2.42	V44	2
V13	2.82	V27	3	V9	NA	V9	NA	V37	2.50	V37	2.5
V43	2.82	V28	3	V17	NA	V17	NA	V38	2.64	V3	3
V5	3.00	V31	3	V35	NA	V35	NA	V5	2.68	V5	3
V20	3.00	V35	3	V36	NA	V36	NA	V35	2.75	V22	3
V22	3.10	V37	3	V37	NA	V37	NA	V22	2.78	V35	3
V4	3.20	V43	3	V38	NA	V38	NA	V3	2.79	V38	3

Importance Rank based on the Department (Con.)				Importance Rank based on Retail Format							
Other Departments				**Food and Beverage Stores**				**Other Retail Formats**			
Rank	Mean	Rank	Median	Rank	Mean	Rank	Median	Rank	Mean	Rank	Median
V30	1.33	V6	1	V2	1.38	V1	1	V9	1.25	V1	1
V18	1.44	V18	1	V32	1.42	V2	1	V1	1.47	V2	1
V19	1.44	V19	1	V19	1.43	V7	1	V32	1.52	V8	1
V32	1.44	V24	1	V1	1.47	V19	1	V33	1.52	V9	1
V1	1.56	V30	1	V33	1.50	V32	1	V2	1.56	V12	1
V2	1.56	V32	1	V30	1.61	V33	1	V12	1.58	V30	1
V6	1.56	V33	1	V7	1.61	V3	2	V30	1.61	V32	1
V7	1.56	V1	2	V18	1.64	V4	2	V6	1.67	V3	2
V10	1.56	V2	2	V44	1.68	V5	2	V8	1.72	V5	2
V24	1.56	V3	2	V29	1.69	V6	2	V10	1.77	V6	2
V33	1.56	V4	2	V9	1.70	V8	2	V18	1.81	V7	2
V41	1.67	V5	2	V6	1.71	V9	2	V21	1.82	V10	2
V44	1.67	V7	2	V10	1.71	V10	2	V19	1.85	V11	2
V9	1.75	V8	2	V8	1.81	V11	2	V31	1.88	V13	2
V29	1.75	V9	2	V36	1.83	V12	2	V14	1.94	V14	2
V14	1.78	V10	2	V14	1.84	V13	2	V17	2.00	V15	2
V23	1.78	V11	2	V24	1.86	V14	2	V37	2.00	V16	2
V8	1.89	V12	2	V28	1.91	V15	2	V44	2.00	V17	2
V11	1.89	V14	2	V11	1.91	V16	2	V20	2.03	V18	2
V12	1.89	V15	2	V21	1.93	V17	2	V11	2.06	V19	2
V21	1.89	V16	2	V41	1.99	V18	2	V16	2.09	V20	2
V28	1.89	V17	2	V40	2.00	V20	2	V13	2.15	V21	2
V3	2.00	V20	2	V16	2.01	V21	2	V41	2.18	V23	2
V4	2.00	V21	2	V31	2.03	V23	2	V15	2.21	V24	2
V15	2.00	V23	2	V12	2.05	V24	2	V28	2.24	V25	2
V17	2.00	V25	2	V39	2.06	V25	2	V36	2.25	V26	2
V36	2.00	V26	2	V15	2.09	V26	2	V26	2.31	V28	2
V38	2.00	V28	2	V23	2.15	V27	2	V39	2.36	V29	2
V39	2.00	V29	2	V17	2.19	V28	2	V5	2.39	V31	2
V40	2.00	V31	2	V34	2.19	V29	2	V24	2.39	V33	2
V20	2.11	V34	2	V38	2.19	V30	2	V43	2.41	V34	2
V31	2.11	V35	2	V42	2.27	V31	2	V40	2.42	V36	2
V42	2.11	V36	2	V25	2.33	V34	2	V7	2.44	V37	2
V16	2.22	V38	2	V4	2.36	V36	2	V25	2.47	V39	2
V34	2.25	V39	2	V26	2.37	V38	2	V27	2.47	V40	2
V5	2.33	V40	2	V3	2.37	V39	2	V29	2.47	V41	2
V26	2.33	V41	2	V20	2.37	V40	2	V3	2.50	V42	2
V35	2.33	V42	2	V5	2.41	V41	2	V4	2.50	V43	2
V43	2.33	V43	2	V27	2.44	V42	2	V38	2.50	V44	2
V25	2.56	V44	2	V13	2.45	V43	2	V34	2.53	V27	2.5
V27	2.56	V13	3	V43	2.45	V44	2	V23	2.55	V38	2.5
V13	2.67	V22	3	V37	2.45	V37	2.5	V42	2.56	V4	3
V22	2.67	V27	3	V35	2.71	V22	3	V35	3.00	V22	3
V37	2.75	V37	3	V22	2.85	V35	3	V22	3.16	V35	3

Importance Rank based on Country											
US				**UK**				**Taiwan**			
Rank	Mean	Rank	Median	Rank	Mean	Rank	Median	Rank	Mean	Rank	Median
V7	1.25	V1	1	V9	1.00	V2	1	V2	1.41	V1	1
V1	1.40	V2	1	V17	1.00	V8	1	V32	1.42	V2	1
V6	1.48	V6	1	V36	1.00	V9	1	V19	1.46	V19	1
V32	1.48	V7	1	V2	1.30	V17	1	V1	1.48	V32	1
V12	1.52	V10	1	V1	1.50	V36	1	V33	1.50	V33	1
V30	1.52	V12	1	V33	1.50	V1	1.5	V30	1.59	V3	2
V33	1.52	V32	1	V8	1.60	V10	1.5	V18	1.65	V5	2
V2	1.55	V33	1	V10	1.60	V30	1.5	V44	1.65	V6	2
V10	1.58	V3	2	V32	1.60	V32	1.5	V7	1.66	V7	2
V18	1.75	V5	2	V6	1.70	V33	1.5	V29	1.69	V8	2
V19	1.76	V8	2	V21	1.70	V4	2	V9	1.69	V9	2
V21	1.76	V11	2	V12	1.80	V5	2	V6	1.74	V10	2
V14	1.81	V13	2	V19	1.80	V6	2	V10	1.76	V11	2
V31	1.86	V14	2	V18	1.90	V11	2	V8	1.77	V12	2
V15	1.90	V15	2	V4	2.00	V12	2	V14	1.82	V13	2
V20	1.95	V16	2	V11	2.00	V13	2	V24	1.85	V14	2
V13	1.95	V18	2	V16	2.00	V14	2	V36	1.86	V15	2
V16	1.95	V19	2	V30	2.00	V15	2	V28	1.89	V16	2
V3	2.00	V20	2	V35	2.00	V16	2	V11	1.91	V17	2
V8	2.00	V21	2	V5	2.10	V18	2	V21	1.95	V18	2
V44	2.05	V24	2	V20	2.10	V19	2	V41	1.98	V20	2
V11	2.10	V25	2	V31	2.10	V20	2	V39	2.01	V21	2
V41	2.10	V26	2	V15	2.20	V21	2	V40	2.01	V23	2
V5	2.24	V28	2	V28	2.20	V24	2	V31	2.02	V24	2
V40	2.24	V29	2	V13	2.30	V28	2	V12	2.03	V25	2
V28	2.38	V30	2	V44	2.30	V31	2	V16	2.04	V26	2
V26	2.40	V31	2	V14	2.50	V34	2	V23	2.08	V27	2
V43	2.45	V34	2	V24	2.50	V35	2	V15	2.15	V28	2
V24	2.48	V40	2	V34	2.50	V43	2	V17	2.19	V29	2
V25	2.50	V41	2	V41	2.50	V44	2	V38	2.20	V30	2
V34	2.50	V43	2	V43	2.50	V23	2.5	V34	2.20	V31	2
V29	2.52	V44	2	V29	2.56	V40	2.5	V42	2.24	V34	2
V39	2.52	V27	2.5	V23	2.70	V41	2.5	V25	2.28	V36	2
V27	2.60	V42	2.5	V39	2.70	V3	3	V26	2.30	V37	2
V42	2.70	V22	3	V42	2.70	V7	3	V4	2.37	V38	2
V22	2.85	V23	3	V3	2.80	V22	3	V20	2.38	V39	2
V23	2.95	V39	3	V40	2.80	V25	3	V27	2.38	V40	2
V4	NA	V4	NA	V26	2.90	V26	3	V37	2.42	V41	2
V9	NA	V9	NA	V27	2.90	V27	3	V3	2.43	V42	2
V17	NA	V17	NA	V7	3.00	V29	3	V43	2.43	V43	2
V35	NA	V35	NA	V25	3.00	V37	3	V5	2.46	V44	2
V36	NA	V36	NA	V37	3.00	V38	3	V13	2.47	V4	2.5
V37	NA	V37	NA	V38	3.00	V39	3	V35	2.73	V22	3
V38	NA	V38	NA	V22	3.10	V42	3	V22	2.91	V35	3

Appendix H: Kolmogorov-Smirnov Test for Department Comparison

MARKETING vs. OPERATIONS AND STORE DEVELOPMENT												
V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.55	1.00	0.12	0.37	1.00	0.27	0.35	0.44	0.85	0.63	0.93	0.11	0.98
MARKETING vs. ACCOUNTING, FINANCE AND AUDITING												
V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.26	0.32	0.05	0.62	0.99	1.00	0.94	0.45	1.00	1.00	0.34	0.59	1.00
MARKETING vs. HUMAN RESOURCES												
V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.52	1.00	0.99	0.93	0.17	0.35	1.00	0.00	0.69	0.13	1.00	1.00	1.00
MARKETING vs. INVESTOR RELATIONS												
V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.55	0.58	0.10	0.77	0.18	0.23	0.27	0.02	1.00	0.35	0.96	0.54	NA
MARKETING vs. R&D AND INFORMATION SYSTEM												
V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.81	0.85	0.39	1.00	1.00	0.99	0.60	0.28	1.00	1.00	1.00	0.60	0.88
MARKETING vs. OTHER DEPARTMENTS												
V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.64	0.97	0.30	1.00	0.99	1.00	0.87	0.99	1.00	0.97	0.98	0.70	1.00
OPERATIONS AND STORE DEVELOPMENT vs. ACCOUNTING, FINANCE AND AUDITING												
V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
1.00	0.59	0.66	0.99	0.74	0.13	0.72	0.38	0.98	0.92	0.76	0.90	0.81
OPERATIONS AND STORE DEVELOPMENT vs. HUMAN RESOURCES												
V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.02	1.00	0.05	0.23	0.09	0.90	0.53	0.13	0.09	0.01	0.98	0.15	1.00
OPERATIONS AND STORE DEVELOPMENT vs. INVESTOR RELATIONS												
V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.85	0.80	0.77	0.85	0.41	0.02	0.49	0.00	0.92	0.25	0.02	0.70	NA
OPERATIONS AND STORE DEVELOPMENT vs. R&D AND INFORMATION SYSTEM												
V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
1.00	0.41	1.00	0.94	0.91	0.28	0.19	1.00	0.95	0.70	1.00	0.78	0.58
OPERATIONS AND STORE DEVELOPMENT vs. OTHER DEPARTMENTS												
V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
1.00	1.00	0.97	1.00	1.00	0.59	1.00	1.00	0.82	0.18	0.97	1.00	0.97

ACCOUNTING, FINANCE AND AUDITING vs. HUMAN RESOURCES

V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.00	0.98	0.02	0.39	0.01	0.60	0.98	0.00	0.30	0.03	0.32	0.60	0.93

ACCOUNTING, FINANCE AND AUDITING vs. INVESTOR RELATIONS

V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
1.00	0.04	0.85	0.65	0.06	0.10	0.06	0.01	1.00	0.00	0.09	0.07	NA

ACCOUNTING, FINANCE AND AUDITING vs. R&D AND INFORMATION SYSTEM

V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.99	0.01	0.56	0.89	1.00	1.00	0.94	0.23	1.00	0.97	0.94	1.00	1.00

ACCOUNTING, FINANCE AND AUDITING vs. OTHER DEPARTMENTS

V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
1.00	1.00	0.28	1.00	1.00	1.00	0.76	1.00	1.00	0.61	0.76	1.00	1.00

HUMAN RESOURCES vs. INVESTOR RELATIONS

V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
1.00	0.25	0.07	0.01	0.86	0.46	0.21	0.55	0.14	1.00	0.94	0.99	NA

HUMAN RESOURCES vs. R&D AND INFORMATION SYSTEM

V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.04	0.62	0.16	0.92	0.12	0.92	0.74	0.33	0.59	0.29	0.80	0.74	0.64

HUMAN RESOURCES vs. OTHER DEPARTMENTS

V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.08	0.99	0.13	0.60	0.06	0.60	0.82	0.16	1.00	0.07	0.60	0.64	1.00

INVESTOR RELATIONS vs. R&D AND INFORMATION SYSTEM

V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
0.04	0.99	0.49	0.84	0.23	0.07	0.94	0.13	1.00	0.10	0.07	0.04	NA

INVESTOR RELATIONS vs. OTHER DEPARTMENTS

V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
1.00	1.00	0.84	1.00	0.11	0.26	0.98	0.20	0.72	0.11	0.68	0.56	NA

R&D AND INFORMATION SYSTEM vs. OTHER DEPARTMENTS

V4	V10	V11	V12	V20	V23	V24	V28	V29	V31	V39	V40	V42
1.00	0.48	1.00	0.96	1.00	1.00	0.29	1.00	1.00	0.95	1.00	1.00	1.00

Appendix I: Survey Descriptive Analysis (Case Study)

Importance Rank based on the Department (Expected Data)											
Marketing				**Operations and Store Development**				**Accounting, Finance and Auditing**			
Rank	Mean	Rank	Median	Rank	Mean	Rank	Median	Rank	Mean	Rank	Median
V19	1.47	V19	1.00	V19	1.24	V1	1.00	V1	1.29	V1	1.00
V32	1.53	V33	1.00	V32	1.24	V2	1.00	V2	1.29	V2	1.00
V33	1.60	V1	2.00	V33	1.24	V7	1.00	V19	1.46	V19	1.00
V1	1.67	V2	2.00	V7	1.35	V9	1.00	V32	1.54	V32	1.00
V2	1.67	V3	2.00	V9	1.41	V19	1.00	V8	1.58	V11	1.50
V44	1.67	V6	2.00	V29	1.41	V29	1.00	V6	1.63	V33	1.50
V28	1.80	V7	2.00	V30	1.47	V30	1.00	V33	1.63	V4	2.00
V36	1.80	V8	2.00	V18	1.63	V32	1.00	V9	1.67	V6	2.00
V7	1.87	V9	2.00	V2	1.65	V33	1.00	V10	1.67	V7	2.00
V8	1.87	V10	2.00	V8	1.65	V40	1.00	V18	1.67	V8	2.00
V18	1.87	V12	2.00	V24	1.65	V18	1.50	V28	1.67	V9	2.00
V29	1.87	V13	2.00	V40	1.65	V3	2.00	V11	1.71	V10	2.00
V30	1.93	V14	2.00	V1	1.71	V4	2.00	V30	1.75	V12	2.00
V9	2.00	V15	2.00	V6	1.75	V6	2.00	V7	1.79	V14	2.00
V10	2.00	V16	2.00	V36	1.76	V8	2.00	V17	1.79	V15	2.00
V23	2.00	V18	2.00	V44	1.76	V10	2.00	V40	1.83	V16	2.00
V24	2.00	V21	2.00	V28	1.88	V11	2.00	V14	1.83	V17	2.00
V14	2.07	V23	2.00	V31	1.88	V12	2.00	V21	1.88	V18	2.00
V6	2.13	V24	2.00	V10	1.88	V13	2.00	V36	1.88	V20	2.00
V31	2.13	V26	2.00	V11	1.94	V14	2.00	V44	1.88	V21	2.00
V41	2.13	V28	2.00	V14	2.00	V15	2.00	V12	1.92	V23	2.00
V26	2.20	V29	2.00	V21	2.00	V16	2.00	V29	1.92	V24	2.00
V21	2.21	V30	2.00	V38	2.00	V17	2.00	V38	1.96	V25	2.00
V12	2.27	V31	2.00	V39	2.06	V20	2.00	V39	1.96	V26	2.00
V15	2.27	V32	2.00	V15	2.06	V21	2.00	V25	2.04	V28	2.00
V16	2.27	V34	2.00	V4	2.12	V23	2.00	V41	2.04	V29	2.00
V3	2.33	V36	2.00	V12	2.12	V24	2.00	V23	2.04	V30	2.00
V39	2.33	V38	2.00	V41	2.12	V25	2.00	V4	2.08	V31	2.00
V13	2.40	V39	2.00	V16	2.18	V26	2.00	V16	2.13	V34	2.00
V40	2.40	V40	2.00	V23	2.18	V27	2.00	V31	2.13	V36	2.00
V5	2.43	V41	2.00	V27	2.18	V28	2.00	V24	2.17	V37	2.00
V38	2.46	V44	2.00	V25	2.24	V31	2.00	V15	2.21	V38	2.00
V34	2.47	V5	2.50	V37	2.24	V34	2.00	V42	2.21	V39	2.00
V17	2.53	V37	2.50	V17	2.29	V36	2.00	V34	2.25	V40	2.00
V27	2.53	V4	3.00	V26	2.35	V37	2.00	V37	2.29	V41	2.00
V42	2.53	V11	3.00	V34	2.41	V38	2.00	V26	2.33	V42	2.00
V37	2.57	V17	3.00	V5	2.44	V39	2.00	V43	2.38	V43	2.00
V20	2.60	V20	3.00	V20	2.47	V41	2.00	V20	2.46	V44	2.00
V25	2.60	V22	3.00	V13	2.50	V44	2.00	V13	2.48	V3	3.00
V35	2.62	V25	3.00	V43	2.53	V5	2.50	V5	2.54	V5	3.00
V4	2.73	V27	3.00	V3	2.56	V22	3.00	V3	2.58	V13	3.00
V43	2.73	V35	3.00	V42	2.76	V35	3.00	V27	2.63	V22	3.00
V11	2.80	V42	3.00	V35	2.87	V42	3.00	V35	2.83	V27	3.00
V22	2.93	V43	3.00	V22	2.94	V43	3.00	V22	3.09	V35	3.00

Importance Rank based on the Department (Expected Data)							
R&D and Information systems				**Other Departments**			
Rank	Mean	Rank	Median	Rank	Mean	Rank	Median
V1	1.36	V1	1.00	V2	1.36	V2	1.00
V2	1.57	V19	1.00	V7	1.62	V1	2.00
V19	1.57	V2	1.50	V1	1.71	V4	2.00
V9	1.64	V32	1.50	V19	1.71	V6	2.00
V18	1.64	V6	2.00	V32	1.71	V7	2.00
V10	1.77	V7	2.00	V8	1.79	V8	2.00
V32	1.79	V8	2.00	V9	1.79	V9	2.00
V8	1.86	V9	2.00	V30	1.79	V10	2.00
V33	1.86	V10	2.00	V44	1.79	V11	2.00
V7	1.93	V12	2.00	V33	1.86	V12	2.00
V30	1.93	V14	2.00	V36	1.86	V14	2.00
V6	2.00	V18	2.00	V18	1.93	V15	2.00
V36	2.00	V21	2.00	V29	1.93	V16	2.00
V44	2.00	V23	2.00	V40	1.93	V17	2.00
V14	2.07	V24	2.00	V6	2.00	V18	2.00
V24	2.07	V29	2.00	V14	2.00	V19	2.00
V38	2.15	V30	2.00	V39	2.00	V20	2.00
V21	2.21	V33	2.00	V11	2.07	V21	2.00
V29	2.21	V34	2.00	V16	2.07	V23	2.00
V41	2.21	V36	2.00	V23	2.07	V24	2.00
V12	2.29	V38	2.00	V24	2.07	V27	2.00
V23	2.29	V39	2.00	V41	2.07	V29	2.00
V40	2.29	V40	2.00	V42	2.07	V30	2.00
V39	2.36	V41	2.00	V17	2.14	V31	2.00
V34	2.38	V44	2.00	V10	2.21	V32	2.00
V42	2.38	V35	2.50	V15	2.21	V33	2.00
V17	2.43	V3	3.00	V4	2.29	V34	2.00
V15	2.46	V4	3.00	V12	2.29	V36	2.00
V35	2.50	V5	3.00	V21	2.36	V39	2.00
V16	2.54	V11	3.00	V28	2.36	V40	2.00
V26	2.54	V13	3.00	V31	2.36	V41	2.00
V28	2.57	V15	3.00	V34	2.36	V42	2.00
V31	2.62	V16	3.00	V13	2.43	V44	2.00
V37	2.64	V17	3.00	V43	2.43	V13	2.50
V25	2.67	V20	3.00	V27	2.46	V25	2.50
V11	2.71	V22	3.00	V25	2.50	V26	2.50
V20	2.71	V25	3.00	V37	2.50	V28	2.50
V43	2.71	V26	3.00	V20	2.57	V37	2.50
V3	2.77	V27	3.00	V26	2.57	V43	2.50
V4	2.86	V28	3.00	V22	2.62	V3	3.00
V13	2.86	V31	3.00	V38	2.64	V5	3.00
V5	2.93	V37	3.00	V35	2.75	V22	3.00
V27	3.08	V42	3.00	V3	2.79	V35	3.00
V22	3.15	V43	3.00	V5	2.86	V38	3.00

Importance Rank based on the Department (Actual Data)											
Marketing				**Operations and Store Development**				**Accounting, Finance and Auditing**			
Rank	Mean	Rank	Median	Rank	Mean	Rank	Median	Rank	Mean	Rank	Median
V25	1.77	V4	2.00	V25	1.47	V25	1.00	V25	1.71	V2	2.00
V26	2.00	V11	2.00	V32	1.53	V1	2.00	V32	1.96	V4	2.00
V32	2.07	V12	2.00	V4	1.65	V2	2.00	V4	2.08	V7	2.00
V36	2.07	V17	2.00	V36	1.71	V4	2.00	V26	2.13	V11	2.00
V4	2.13	V18	2.00	V26	1.76	V7	2.00	V44	2.13	V12	2.00
V30	2.20	V25	2.00	V12	1.82	V8	2.00	V29	2.21	V17	2.00
V12	2.33	V26	2.00	V33	1.82	V11	2.00	V36	2.21	V19	2.00
V17	2.33	V30	2.00	V24	2.00	V12	2.00	V2	2.33	V20	2.00
V18	2.33	V32	2.00	V1	2.06	V13	2.00	V11	2.38	V21	2.00
V44	2.40	V33	2.00	V30	2.06	V18	2.00	V17	2.38	V25	2.00
V19	2.47	V36	2.00	V44	2.06	V19	2.00	V19	2.38	V26	2.00
V24	2.47	V44	2.00	V2	2.12	V21	2.00	V20	2.38	V29	2.00
V33	2.47	V21	2.50	V19	2.12	V23	2.00	V30	2.42	V32	2.00
V11	2.50	V34	2.50	V40	2.18	V24	2.00	V7	2.43	V36	2.00
V2	2.53	V38	2.50	V7	2.24	V26	2.00	V41	2.43	V41	2.00
V8	2.57	V41	2.50	V21	2.24	V28	2.00	V1	2.46	V44	2.00
V21	2.57	V1	3.00	V29	2.24	V29	2.00	V12	2.46	V1	2.50
V38	2.57	V2	3.00	V31	2.24	V30	2.00	V21	2.46	V24	2.50
V41	2.57	V3	3.00	V8	2.29	V31	2.00	V24	2.46	V30	2.50
V9	2.60	V5	3.00	V9	2.29	V32	2.00	V31	2.50	V3	3.00
V29	2.60	V6	3.00	V28	2.29	V33	2.00	V33	2.50	V5	3.00
V22	2.64	V7	3.00	V38	2.29	V34	2.00	V18	2.54	V6	3.00
V34	2.64	V8	3.00	V23	2.35	V36	2.00	V15	2.57	V8	3.00
V20	2.67	V9	3.00	V11	2.38	V38	2.00	V8	2.58	V9	3.00
V7	2.71	V10	3.00	V18	2.41	V40	2.00	V9	2.58	V10	3.00
V10	2.73	V13	3.00	V20	2.41	V44	2.00	V28	2.59	V13	3.00
V13	2.73	V14	3.00	V34	2.44	V5	2.50	V38	2.61	V14	3.00
V23	2.73	V15	3.00	V14	2.47	V15	2.50	V10	2.63	V15	3.00
V28	2.73	V16	3.00	V5	2.50	V3	3.00	V40	2.63	V16	3.00
V31	2.73	V19	3.00	V10	2.53	V6	3.00	V39	2.65	V18	3.00
V42	2.79	V20	3.00	V13	2.53	V9	3.00	V34	2.68	V22	3.00
V39	2.86	V22	3.00	V17	2.53	V10	3.00	V23	2.70	V23	3.00
V3	2.87	V23	3.00	V15	2.56	V14	3.00	V42	2.74	V27	3.00
V40	2.92	V24	3.00	V41	2.65	V16	3.00	V14	2.78	V28	3.00
V1	2.93	V27	3.00	V22	2.73	V17	3.00	V27	2.78	V31	3.00
V14	3.00	V28	3.00	V3	2.76	V20	3.00	V6	2.83	V33	3.00
V43	3.00	V29	3.00	V27	2.76	V22	3.00	V37	2.86	V34	3.00
V5	3.07	V31	3.00	V39	2.76	V27	3.00	V43	2.86	V35	3.00
V27	3.07	V37	3.00	V16	2.81	V35	3.00	V13	2.87	V37	3.00
V37	3.08	V39	3.00	V6	2.82	V37	3.00	V3	2.88	V38	3.00
V15	3.13	V40	3.00	V37	2.82	V39	3.00	V22	2.90	V39	3.00
V16	3.13	V42	3.00	V43	2.86	V41	3.00	V5	2.96	V40	3.00
V6	3.27	V43	3.00	V42	2.87	V42	3.00	V16	3.09	V42	3.00
V35	3.36	V35	4.00	V35	3.00	V43	3.00	V35	3.15	V43	3.00

Importance Rank based on the Department (Actual Data)							
R&D and Information systems				**Other Departments**			
Rank	Mean	Rank	Median	Rank	Mean	Rank	Median
V25	1.71	V1	2.00	V4	1.64	V1	2.00
V4	2.07	V2	2.00	V25	1.71	V2	2.00
V26	2.08	V4	2.00	V36	1.79	V4	2.00
V18	2.14	V11	2.00	V44	1.86	V8	2.00
V2	2.21	V12	2.00	V29	2.07	V9	2.00
V30	2.21	V17	2.00	V18	2.14	V10	2.00
V12	2.31	V18	2.00	V21	2.14	V11	2.00
V36	2.31	V25	2.00	V26	2.15	V12	2.00
V44	2.31	V26	2.00	V30	2.21	V13	2.00
V28	2.33	V28	2.00	V2	2.29	V17	2.00
V11	2.36	V30	2.00	V8	2.29	V18	2.00
V17	2.36	V32	2.00	V9	2.29	V19	2.00
V32	2.36	V36	2.00	V10	2.29	V20	2.00
V1	2.43	V38	2.00	V13	2.29	V21	2.00
V19	2.43	V41	2.00	V19	2.29	V23	2.00
V9	2.50	V44	2.00	V23	2.29	V24	2.00
V24	2.50	V9	2.50	V39	2.29	V25	2.00
V33	2.50	V13	2.50	V41	2.29	V26	2.00
V38	2.50	V14	2.50	V1	2.36	V28	2.00
V40	2.54	V19	2.50	V11	2.36	V29	2.00
V41	2.54	V21	2.50	V12	2.36	V30	2.00
V13	2.57	V24	2.50	V14	2.36	V36	2.00
V14	2.57	V3	3.00	V20	2.36	V39	2.00
V20	2.57	V5	3.00	V28	2.36	V40	2.00
V29	2.57	V6	3.00	V32	2.36	V41	2.00
V23	2.62	V7	3.00	V17	2.43	V44	2.00
V21	2.64	V8	3.00	V24	2.43	V14	2.50
V39	2.67	V10	3.00	V40	2.43	V32	2.50
V7	2.69	V15	3.00	V31	2.50	V33	2.50
V22	2.69	V16	3.00	V7	2.54	V3	3.00
V8	2.71	V20	3.00	V33	2.57	V5	3.00
V10	2.71	V22	3.00	V34	2.57	V6	3.00
V31	2.71	V23	3.00	V42	2.57	V7	3.00
V34	2.73	V27	3.00	V27	2.58	V15	3.00
V42	2.75	V29	3.00	V22	2.69	V16	3.00
V27	2.77	V31	3.00	V37	2.79	V22	3.00
V37	2.80	V33	3.00	V5	2.93	V27	3.00
V15	2.93	V34	3.00	V6	2.93	V31	3.00
V3	3.00	V35	3.00	V15	2.93	V34	3.00
V5	3.00	V37	3.00	V43	2.93	V35	3.00
V6	3.00	V39	3.00	V3	3.00	V37	3.00
V35	3.00	V40	3.00	V16	3.07	V38	3.00
V43	3.08	V42	3.00	V38	3.07	V42	3.00
V16	3.14	V43	3.00	V35	3.27	V43	3.00

<u>Paper One:</u>

Measuring Retail Company Performance Using Credit Scoring Techniques

- Submitted to *European Journal of Operational Research* (Under Review)
- Presented in the Proceedings of *<u>Credit Scoring and Credit Control IX</u>*, Credit Research Centre, Management School, University of Edinburgh, Sep. 7- Sep. 9. 2005
- Presented at *<u>Taiwan Scientific Symposium in Cambridge</u>*, CTS, SUROCUK and Science and Technology Division, Taipei Representative Office in the UK, Trinity College, University of Cambridge, Dec. 5 2005
- Presented in the Proceedings of *<u>Paris Finance International Meeting</u>*, French Finance Association (AFFI), BNP-PARIBAS, Paris, Dec. 19 – Dec. 20, 2005
- Presented in the Proceedings of *<u>Campus for Finance Research Conference: Fixed Income, Lending, Borrowing and Taking Risk</u>*, WHU Campus for Finance, Vallendar, Germany, Jan. 10 – Jan. 12, 2006
- Presented in the Proceedings of *<u>2006 International Conference in Business and Information (BAI)</u>*, Singapore, July 12 – July 14

O.R. Applications

Measuring Retail Company Performance Using Credit Scoring Techniques

Yu-Chiang Hu[a,1] and Jake Ansell[a]

[a]Management School and Economics, University of Edinburgh, William Robertson Building, 50 George Square, Edinburgh, EH8 9JY, UK

Abstract

This paper proposes a theoretical framework for predicting financial distress based on Hunt's (2000) Resource-Advantage Theory of Competition. The study focuses on the US retail market. Five credit scoring methodologies: Naïve Bayes, Logistic Regression, Recursive Partitioning, Artificial Neural Network, and Sequential Minimal Optimization (SMO), are used on a sample of 195 healthy companies and 51 distressed firms over five time periods from 1994 to 2002.

Analyses provide sufficient evidence that the five credit scoring methodologies have sound classification ability in the time period of one year before financial distress. Moreover, the methodologies remain sound even five years prior to financial distress with classification accuracy rates above 80% and AUROC values above 0.80. However, it is difficult to conclude that which modelling methodology has the absolute best classification ability, since the model's performance varies in terms of different time scales and different variable groups.

This paper also shows external environment influences exist based on all five credit scoring models, but these influences are weak. With regards to the model applicability, a subset of the different models is compared with Moody's rankings. It is found that both SMO and logistic regression models are better than the neural network model in terms of similarity with Moody's ranking, with logistic regression model being slightly better than the SMO Model.

Keywords: Finance; Credit scoring; Retailing; Multivariate statistics; Artificial intelligence

1. Introduction

There is considerable effort devoted to the performance measurement of companies and being able to forecast their financial distress. The approaches used have covered a wide range of methodologies, for example, Beaver's (1966) univariate analysis model, Altman's (1968) Z-score model and Ohlson's (1980) logistic regression model.

[1] Corresponding author

Email address: Y.A.Hu@sms.ed.ac.uk (Yu-Chiang Hu), Jake.Ansell@ed.ac.uk (Jake Ansell)

It has been argued by a number of authors that generic models for all sectors tend to be too general and lack the ability to deal with specific industrial sectors. It was decided to focus on the retail sector, since according to Dawson (2000) retail risk assessment and evaluation will be a critical area of research. The USA retail sector was chosen because of the clear definitions and reporting of financial distress through Chapters 7 and 11. A sample of 195 healthy companies and 51 distressed firms were selected from 1994 to 2002. Timescale is clearly an issue and in the paper this is explored, the results unsurprisingly find that, for most models, the year before financial distress provides the best prediction, though, up to at least 5 years provide good prediction.

A range of variables that can be assembled to describe the performance of retails companies. In the current research, 170 potential performance measures have been considered which cover both internal and external measures, based on Resource-Advantage theory (Hunt, 2000). Yet obviously with such a large number of variables to choose from there is a danger of over-fitting and so there is a need to limit the number of variables. After exploring a range of models and taking into account the sample size, five variables were employed in the final analysis.

In this paper, five credit scoring methodologies are used: Naïve Bayes, Logistic Regression, Recursive Partitioning, Artificial Neural Network, and Sequential Minimal Optimization (SMO). These models were fitted to the data and the all performed well. Since the size of dataset did not allow a hold out sample, it was felt that a comparison should be made with an alternative external rating, and Moody's rating was chosen. The results indicated that the most comparable models were the logistic regression model and the SMO model.

The next section will discuss the alternative credit scoring modelling approaches considered. Section 3 will initially discuss the measures that could be used to determine financial distress, and then proceed to describe those variables that have been used in the study. Details of the sample selected will be given at the end of the section. Section 4 will describe the approach taken in fitting the models. This is followed by the results of the analysis. Finally there will be a discussion of the results in section 6.

2. Credit Scoring Modelling

Beaver (1966) was a pioneer in financial distress prediction research with a number of authors following his work. He conducted an analysis of likelihood ratios based on a

Bayesian approach. He argued that the default prediction problem could be regarded as a problem of evaluating the probability of financial distress conditional upon the value of a specific financial ratio. Naïve Bayesian approach provides a simple method to deal with a classification problem. Let *H* be the healthy samples and let *D* be the distressed samples. Moreover, let *X* be a vector of independent variables and let *x* represent a particular vector of an independent variable. The conditional probability of a financial distress company in terms of a specific financial ratio *x* can be expressed as:

$$P(D \mid X = x) = \frac{P(D)P(X = x \mid D)}{P(X = x)} \tag{1}$$

Beaver (1966) used only a single measure and so was limited, but the approach can be easily generalised. Altman (1968) suggested the Multiple Discriminant Analysis (MDA) to develop a Z-score bankruptcy prediction model based on five financial ratios. After Altman's (1968) research, a number of studies also use MDA to predict firm's default, including Deakin (1972), Blum (1974), Libby (1975), Altman et al. (1977), and Taffler (1984). Generally the choice was to fit a Linear Discriminant Function (LDF) based on the normal distribution with the assumption of equal covariance matrices. These assumptions are too restrictive for industrial data, (Eisenbeis, 1977). Deakin (1976) contended that even if after transformation financial ratio data do not follow normal distribution. Moreover, Hamer (1983) evaluated the sensitivity of financial distress prediction models in terms of four different variable sets from previous studies (Altman, 1968; Deakin, 1972; Blum, 1974; Ohlson, 1980). She pointed out that the covariance matrices in each variable set were statistically different.

Ohlson (1980) was the first to apply the conditional probability model and in particular, the logistic model, to bankruptcy prediction research. Unlike MDA, the logistic model does not require multivariate normality or the equality of covariance matrices of two populations. By logit transformation on odd ratio function, the logistic model can be linearized and used to solve classification problems. A logistic regression can be expressed as follow:

$$\begin{aligned} g(x) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \\ &= \beta \times x^T \end{aligned} \tag{2}$$

where $\pi(x)$ is the logistic function,

$$\pi(x) = \frac{1}{1 + e^{-(\beta \times x^T)}} = \frac{e^{\beta \times x^T}}{1 + e^{\beta \times x^T}} \tag{3}$$

Following Ohlson's (1980) study, Mensah (1983), Casey and Bartczak (1985), as well as Gentry et al. (1985), also employed the conditional probability models to predict financial distress.

In the mid-1980s, Recursive Partitioning Analysis (RPA) or Decision Tree was introduced in the financial distress prediction research area. (Frydman et al., 1985; Marais et al., 1984; Carson and Hoyt, 1995) RPA is a non-parametric technique and does not suffer the limitations from MDA or logistic model. Although Fisher's (1936) linear discriminant method is often viewed as the oldest classification technique, Hand (1997) argued that the basic idea of RPA is very straightforward, and hence the oldest conceptually.

RPA can be regarded as a stepwise procedure. The first step is to select an independent variable as the best discriminator and to decide a cutpoint based on the lowest expected misclassification cost. Based on the cutpoint, the second step is to divide both healthy and distressed firms into two sub-nodes. The third step is to select another (or the same) discriminator and further partition the healthy and distressed firms into another two sub-nodes. The same process can be continued, if further splitting is necessary. Thomas et al. (2002) mentioned two reasons to stop the partitioning process. First, if the number of samples in a node is too small, then further partition is not appropriate. Second, if the classification results between the old node and new nodes do not have significant differences, then it is also not necessary to split the old node. One of the major problems relative to RPA is overfitting: the continuous partitioning process is likely to encourage miss-classification in the terminal node. The overfitting problem can be overcome by a Cross-Validation procedure.

From the late 1980s, the Artificial Intelligence (AI) or Machine Learning Techniques, such as Artificial Neural Networks (ANN), were successfully applied to financial distress prediction studies. A large number of studies compared ANN's prediction performance with other classification methods and proved that ANN had better prediction performance than other methods, (Coats and Fant, 1993; Zhang et al., 1999). The most popular ANN algorithm in the financial distress prediction domain is the Multilayer Perceptron (MLP). A MLP has three main components: input layer, hidden layer and output layer.

The input layer is responsible for receiving information from the outside environment and transferring it to the hidden layer. In the hidden layer, a neuron will assign a series of weights to the inputs, cope with the information via a training process, and then forward the results with weights to the output layer. The training process can be regarded as a weighting determination process. The most frequently used algorithm for training process is the Back Propagation Algorithm (BPA).

Thomas et al. (2002) pointed out that BPA first calculates the difference between the expected output value and the observed output value (called *error*) in the output layer. The next step is to distribute the error back to the network in terms of a weight and to adjust the weight to decrease the error. The process is repeated for all cases, called an *epoch*. After several epochs training, the learning error will reduce to a minimum level and the training process ends. Trigueiros and Taffler (1996) mentioned some advantages of MLP, such as the independence from statistical distribution assumptions. However, MLP also has some limitations. For example, it does not provide adequate significance tests and requires considerable computer power and skills (Tam and Kiang, 1992).

In the late 1990s, another machine learning technique, Support Vector Machine (SVM), was introduced to deal with the classification problem. Fan and Palaniswami (2000) applied SVM to select the financial distress predictors. They pointed out that SVM created an optimal separating hyperplane in the hidden feature space in terms of the principle of structure risk minimization and used the quadratic programming to obtain an optimal solution. However, Platt (1999) argued that a large volume of quadratic programming in SVM training is time consuming. As a result, he introduced a new algorithm, Sequential Minimal Optimization (SMO), to improve the SVM training time. Unlike the previous SVM training methods, SMO only uses two Lagrange multipliers at each training step. It was found that SMO has better performance than other SVM training methods in terms of many aspects, such as better scaling with training sample size.

Some other methodologies were also applied to the financial distress prediction research area and have shown a good performance, including the Rough Sets Approach (Dimitras et al., 1999; McKee and Lensberg, 2002) and the Multidimensional Scaling Approach (Mar-Molinero and Serrano-Cinca, 2001).

3. Performance Measures Selection and Data Collection

3.1 Previous Research Survey

Most of the academic literature has based on the quantitative financial ratios to predict financial distress. However, credit-rating companies including Moody's, S&P, and Fitch take into account both quantitative and qualitative factors with emphasis on the qualitative factors (Moody's, 1998 and 2002; Fitch, 2000 and 2001; S&P, 2002 and 2003). In this paper a large range of measures are explored. These include measured on industrial sector since many authors, (Williams and Goodman, 1971; Gupta and Huefner, 1972; Bowen et al., 1982; Mensah, 1984), have suggested applying the same variables across different sector produces overly general models that overlook the specific attributes of the sectors. Platt and Platt (1990) used industry-related measures in a bankruptcy model and proved that these industry-relative measures could improve the accuracy of the classification model.

In addition, macro-economical factors have significant impact on financial distress prediction models (Rose et al., 1982), and different macroeconomical environments may affect the accuracy of the bankruptcy predictive model, (Mensah, 1984). Other authors have suggested a company's sustainability must be based on cash flow, rather than on earnings in the accounting statements, for earnings include non-cash items that cannot reflect a company's ability to pay back interests or principal, (S&P, 2003). Gentry et al. (1985) developed a financial distress prediction model in terms of a cash flow structure. Although their model showed that only one variable, dividends/cash flow, had significant influence to the bankruptcy prediction.

Finally, the lack of theoretical groundwork for variable selection is a common situation in most financial distress prediction studies. Often, financial distress researchers select independent variables for model construction based on the successful prediction performance in previous studies. Obviously, such a variable selection method is limited and fails to provide a holistic framework for research in financial distress. In contrast, the present research develops a theoretical framework based on Hunt's (2000) Resource-Advantage (R-A) Theory of Competition.

Unlike the traditional perfect competition theory which focuses on factors of production, the R-A theory includes significant qualitative issues such as entrepreneurship and a company's relationship with its suppliers. The theory holds that demand is not only heterogeneous across industries but within them. It also holds that information is imperfect and costly and so that maximising profit is not a viable proposition, one can only seek superior financial performance. Given that companies' resources are different and imperfectly mobile, then Hunt and Morgan (1997) argued that a comparative advantage in resources provides also a comparative advantage in the market place and hence a superior financial performance. The theory suggests seven categories of measures, see table 1.

Table 1 Internal Resources

Internal Resources	**Examples**
Financial Resource	Cash reserves and access to financial markets
Physical Resource	Plant, raw materials, and equipment
Legal Resource	Trademarks and licenses
Human Resource	The skills and knowledge of individual employees and the entrepreneurial skills
Organizational Resource	Controls, routines, cultures, and competences for entrepreneurship
Informational Resource	Knowledge about the market segment, competitors and technology
Relational Resource	Relationships with competitors, suppliers and customer

Source: Modified from Hunt (2000) A General Theory of Competition, Thousand Oaks: Sage pp.128

Based on previous literature survey and interviews with outside stakeholders, 170 potential retail performance measures were obtained. These variables were then studied and classified as Quantifiable Measure - Available Data, Quantifiable Measure - No Available Data, and Difficult to Quantify. Obviously, the analysis focussed on the 67 performance measures in the category Quantifiable Measure - Available Data. These are combined into two groups: Internal Resources Group (G1), and External Factors Group (G2) and are presented in Appendix A. In order to detect external influences, these factors will be re-grouped as G1 and G12 (G1+G2).

3.2 Sample Selection and Data Collection

In this research, data were collected from four main sources: (i) Accounting and Finance Databases, such as, DataStream, Thomson One Banker and OSIRIS, (ii) Annual Report from each sample company, (iii) Government Publications, such as, Budget of the United States Government and (iv) Other sources, such as documents from Organisation for Economic Co-operation and Development (OECD).

In connection with the sample selection of non-defaulting companies, five criteria were considered. Only publicly listed companies were chosen. Given that listed companies had to abide by regulations in the financial market, their financial information tended to be more open and transparent than that of private companies. In addition, small companies were included based on the SBA size standards[2]. This is an improvement from previous studies using the Wall Street Journal Index. The data source is likely to exclude small companies despite the fact that small companies are likely to face financial distress.

[2] SBA's size standards define whether a business entity is small and, thus, eligible for Government programs and preferences reserved for "small business" concerns. Size standards have been established for types of economic activity, or industry, generally under the North American Industry Classification System (NAICS). Information available at: http://ecfr.gpoaccess.gov/

Although Edmister (1972) argued that new firms had great probability of facing financial distress and should be considered in any bankruptcy prediction model, in the present study, only those public sample companies that had been listed for at least three years were considered. There are two reasons to support this criterion. First, a newly listed company may not a new company. Second, studies show that newly listed stocks have abnormal returns after the public announcements of listing, (Sanger and McConnell, 1986). In order to avoid the influences from the newly listed companies, especially for some market relevant measures, no healthy company listed after December 2000 is included. Furthermore, this research does not consider e-retailers, as their performance measures are different from those of traditional retailers. Finally, even if a sample company satisfied the previous four criteria, it is excluded if its data is not complete. As a result of applying the five criteria above, 67 different retail performance measures are collected from a dataset of 195 non-defaulting US retail companies over the time period of 1998 to 2002.

The USA retail sector was chosen because of the clear definitions and reporting of financial distress through Chapters 7 and 11. Based on the US federal bankruptcy law, a financially distressed company might use the bankruptcy code of Chapter 11 to reorganize its financial structure and try to recover from distress, or that of Chapter 7, to go into liquidation and stop all business operations. Drawing on this insight, any company filing for the bankruptcy code of Chapter 11 or Chapter 7, were deemed to be under financial distress and selected for the research.

An important issue is the timing of failed firms' data. Ohlson (1980) suggested that the financial statements prior to the financial distress year should be viewed as the last report, since reports *after* financial distress would usually include the adjustments from auditors in light of the bankruptcy filing. Adopting Ohlson's (1980) viewpoints, data prior to the financial distress year was considered as the last report. Overall, data were collected from 51 financially distressed firms and these companies were divided into five groups in terms of different time scales. (see table 2)

Table 2 Descriptions of Time Scales of Distressed Firms' Data

Group	Number of Failed Firms	Financial Distress Year	Data Collection Time Scale
A	5	2003	From 1998 to 2002
B	13	2002	From 1997 to 2001
C	15	2001	From 1996 to 2000
D	12	2000	From 1995 to 1999
E	6	1999	From 1994 to 1998
Total	51		

4. Methodology

As with any data analysis there need to clean the data to remove outliers. This was done using standard approaches (10-means Cluster Analysis) and reduced the samples by about 5%. Given large number of variables, 67, for consideration would tend towards overfitting. Prior to model construction, a cross-validation process is performed to resolve overfitting. Moore (2001) compared three cross-validation methods: the test set method, the leave-one-out method and the 10-folders method. He argued that the 10-folders cross-validation process only wasted 10% of total data and the training cost was much lower than the leave one out method. Drawing on this insight, the 10-folder method is selected for cross-validation.

Five credit scoring methodologies are employed for model construction: Naïve Bayes, Logistic Regression, Recursive Partitioning, Artificial Neural Network and Sequential Minimal Optimization (SMO). An initial interest of the study was the timescale effect, whether on should use data just prior to the potential financial distress or some time before. Hence, a series of models were fitted to M1 to M5 to allow evaluation of prediction performance from one to five years before financial distress.

Selection of the variables was via two stage model. Hosmer and Lemeshow (2000) suggest that one should initially use univariate analysis to identify the potential variables for the modelling using a *p-value* of 0.25. This was followed by use of forward stepwise model for each approach. The top five variables with higher appearance frequency in each variable group are selected for final model construction, as can be seen in table 3. The Gearing Ratio (V24), Total Debt / (Total Debt + Market Capitalization) (V28) and Operation Cash Flow (V36) are significant variables, since they are common across the models.

Table 3 Stepwise Variable Selection Results

Variable Group	Key Performance Measures
G1	V6: Net Profit Margin V24: Gearing Ratio V28: Total Debt / (Total Debt + Market Capitalization) V36: Operation Cash Flow V55: Payables Turnover
G12	V24: Gearing Ratio V28: Total Debt / (Total Debt + Market Capitalization) V32: Total Assets V36: Operation Cash Flow V56: Five Years Correlation Coefficient between Government Debt / GDP and Total Sales

Model performance was evaluated in terms of two approaches: the Classification Accuracy Rate approach, see Hand (1997), and the Area under the Receiver Operating Characteristics Curve (AUROC) approach, see Thomas et al (2002). In this research, AUROC is applied to the naïve bayes, logistic regression and artificial neural network models. Both the accuracy rate and AUROC are employed for subsequent analyses.

Given the size of the sample available for study it was not possible, and probably it would not have been informative, to employ a hold out sample. Hence, the above methodology will result in potentially overly optimistic results. To overcome this problem for the best modelling approaches, it was decided to compare the results from the study with a standard rating system; in this case Moody's rating. In retailing, there are only 8 rating grades given Aa to C in Moody's system. Hence, the data was ranked according to score and divided into 8 groups. Unfortunately, Moody's ratings were only available for a limited number of companies, since firms undergo the credit rating process due to special circumstances, such as issuing corporate bond. Therefore, the sample size for comparative analysis varies year on year. Logistic regression, neural network and SMO models are selected for the ranking comparison analysis. Again, a range of measures for comparison were used, Kolmogorov-Smirnov (K-S) test, Distance analysis, and Weighted Kappa analysis and finally Graphical Bubble charts.

5. Empirical Analysis

5.1 Time Scale Analysis

As mentioned in section 4, a five-year time scale analysis can be carried out in this research by comparing the performance of models from five different time periods (M1, M2, M3, M4 and M5). M1 is designed for evaluating a model's performance one year before financial distress; M2 is designed for assessing a model's utility two years before the financial distress, and so on. An arrangement of accuracy rate and AUROC results in terms of the five models are expressed in table 4.

Table 4 shows that regardless of the groups of performance measure, M1 has the best classification performance. In addition, even if the time period is five years prior to financial distress, the accuracy rates are above 80% and the AUROC values are above 0.80 among all five modelling methodologies. The results suggest that the overall performance of these five modelling methodologies is sound, even if the time period chosen is as long as five years before financial distress. Furthermore, these results also prove that the key variables selected are effective to predict financial distress.

When comparing the performance of different methodologies, in the G1 variable group, logistic regression model proves to have the best performance one year before the financial distress based on AUROC value and neural network model shows to have best performance one year before the financial distress based on accuracy rate. However, the same cannot be concluded for different variable group. In the G12 variable group, naïve bayes model shows the best performance in terms of the AUROC value and recursive partitioning model presents the best performance based on the accuracy rate one year before financial distress. Moreover, even if in the same variable group, different models show different performance in terms of different time periods.

Furthermore, the same result can also be obtained based on the five years average performance. For example, neural network model shows the best performance in terms of the accuracy rate and the AUROC value in the G1 variable group. However, the same conclusion cannot be achieved, if the model performance is evaluated in the G12 variable group. Drawing on this insight, it is difficult to conclude which modelling methodology has the *absolute* best performance in time scale comparison analysis.

Table 4 Model Performance Evaluation

G1 (Internal Resources Group)							
Methodology	**Performance Measures**	**M1**	**M2**	**M3**	**M4**	**M5**	**Average**
Naïve Bayes	Accuracy Rate (%)	89.02	84.55	80.89	81.30	81.30	***83.412***
	AUROC	0.9161	0.8792	0.8155	0.7798	0.8140	***0.84092***
Logistic Model	Accuracy Rate (%)	89.84	86.99	81.71	82.11	80.89	***84.308***
	AUROC	0.9341	0.8860	0.8156	0.7816	0.7955	***0.84256***
Neural Network	Accuracy Rate (%)	93.09	91.06	87.40	82.93	86.99	***88.294***
	AUROC	0.9158	0.9024	0.8498	0.7982	0.8755	***0.86834***
SMO	Accuracy Rate (%)	89.84	89.02	86.18	82.11	78.46	***85.122***
Recursive Partitioning	Accuracy Rate (%)	92.28	88.21	89.02	85.77	85.77	***88.21***
G12 (Internal Resources Group plus External Factors Group)							
Methodology	**Performance Measures**	**M1**	**M2**	**M3**	**M4**	**M5**	**Average**
Naïve Bayes	Accuracy Rate (%)	91.06	88.62	87.80	86.99	88.62	***88.618***
	AUROC	0.9509	0.9174	0.8967	0.8950	0.9158	***0.91516***
Logistic Model	Accuracy Rate (%)	91.87	89.43	88.21	86.59	87.40	***88.7***
	AUROC	0.9448	0.8970	0.8894	0.8964	0.9079	***0.9071***
Neural Network	Accuracy Rate (%)	90.24	89.43	87.80	87.80	88.21	***88.696***
	AUROC	0.9350	0.9140	0.8762	0.8808	0.8794	***0.89708***
SMO	Accuracy Rate (%)	90.24	89.43	85.77	85.77	87.40	***87.722***
Recursive Partitioning	Accuracy Rate (%)	92.68	89.43	88.21	88.21	89.02	***89.51***

5.2 External Influences Detection Analysis

As mentioned in section 3, the external influences can be detected by comparing the performance of G1 and G12 models. If G12 performs better than G1, external factors have significant impacts on the model classification ability.

In table 4, all G12 models have better classification ability than G1 models founded on the five years average accuracy rate and the five years average AUROC value. However, the performance differences among these models are small. For example, the difference of the average accuracy rate is below 6% and the difference between the average AUROC values is below 0.07 for all models. As a result, it can be concluded that external environment influences exist based on all modelling methodologies, but these influences are weak.

Based on the findings above, the G12 models in the time period of one year before financial distress show the best performance. Results show that regardless the modelling methodologies, the accuracy rates are above 90% and AUROC values are above 0.93. Due to the limitation of sample size, it is impossible to employ a holdout sample, and hence, the current results are potentially overly optimistic. In order to overcome this problem, logistic regression, neural network and SMO models in the time period one year before financial distress in the G12 variable group are selected for the purpose of ranking comparison analysis with Moody's credit rating results.

5.3 Test of Significance

The Kolmogorov-Smirnov test assesses whether two datasets differ significantly. Results of the Kolmogorov-Smirnov Two-Sample test are shown in table 5.

Table 5 Two-Sample Kolmogorov-Smirnov (K-S) test

Modelling Methodology	K-S	2002	2001	2000	1999	1998
Logistic Model	Z Value	1.167	0.993	1.551	1.612	1.241
	p-value	0.131	0.277	0.016	0.011	0.092
Neural Network	Z Value	2.583	2.897	2.041	1.934	1.903
	p-value	0	0	0	0.001	0.001
SMO	Z Value	1.083	1.407	1.551	1.289	1.324
	p-value	0.191	0.038	0.016	0.072	0.06

The highlighted *p*-values in table 5 are not significant at 5% level of significance and indicate when a proposed model provides rankings similar to Moody's. SMO model has similar rankings in years 1998, 1999, and 2002 and logistic regression model has similar results in years 1998, 2001 and 2002. However, neural network does not present any similar ranking result from 1998 to 2002. Significance testing helps determine whether or not there is similarity in ranking. The following techniques attempt to assess the degree of similarity.

5.4 Distance Analysis

The most straightforward approach to compare the degree of similarity between two ordinal data sets is distance analysis. The smaller the distance between the rankings from Moody's and the present study, the better the practical applicability of the study's proposed model. To calculate distances, each cell in the crosstabulation is presented as a proportion of the total sample size. (This allows for year on year comparison, as the sample size of each year is different.) The cell value is then multiplied by the value in the distance matrix. Finally, the resulting values are summed. This gives an overall distance between Moody's model and each of the proposed models. Results are shown in table 6.

Table 6 Overall Distances for Each Modelling Methodology

Modelling Methodology	**2002**	**2001**	**2000**	**1999**	**1998**	***Average Distance***
Logistic Model	0.9861	1.0685	1.2267	1.3117	1.2877	***1.1761***
Neural Network	1.5694	1.7397	1.6133	1.5844	1.3699	***1.5753***
SMO	1.0278	1.3014	1.3867	1.3896	1.3288	***1.2869***

Amongst the three models, the neural network model has the highest average distance between 1998 and 2002, and the highest distances each year. The best model is logistic regression model based on average distance over the five years. The SMO model has similar performance to logistic regression model, and although the average distance is higher.

5.5 Measure of Agreement

Weighted Kappa can be used to measure the concordance between two raters and it is an extension of Cohen's Kappa (1960) suitable for ordinal data and for measuring relative concordance. The values of weighted Kappa are shown in table 7.

Table 7 Weighted Kappa Analysis

Modelling Methodology	**2002**	**2001**	**2000**	**1999**	**1998**	***Average Weighted Kappa***
Logistic Model	0.4368	0.3981	0.3832	0.3714	0.4068	***0.3993***
Neural Network	0.2499	0.2164	0.2874	0.3553	0.4264	***0.3071***
SMO	0.4262	0.3364	0.3575	0.3691	0.4255	***0.3829***

As with distance analysis, average weighted Kappa results suggest that logistic regression model shows the best performance among three models and SMO model has similar performance with logistic regression model. Neural network model still shows the lowest performance in terms of agreement with Moody's.

5.6 Bubble Chart Analysis

In this research, graphical analysis using the bubble chart was developed to facilitate interpretation of similarity. The bubble chart enables a visualization of crosstabulation tables with clear localization of frequencies and a graphical representation of the observations through bubble size.

Bubble charts are interpreted as follows: The closer the bubbles are to the diagonal line, the more similar the rankings are, if the bubbles that are close to the diagonal line are large in size, then it can be concluded that the degree of similarity between rankings is higher and if the bubbles are gathered in the upper left hand corner and in the lower right hand corner, then the degree of similarity between the compared rankings is low.

Based on the distance and measure of agreement analyses, logistic regression model shows the highest degree of similarity with Moody's ratings in the year 2002, (see tables 6 and 7). In contrast, the neural network model presents the worst performance in the comparative study in the year 2001. In order to illustrate the utility of bubble chart analysis, these two counterexamples are presented in figure 1 and figure 2, respectively.

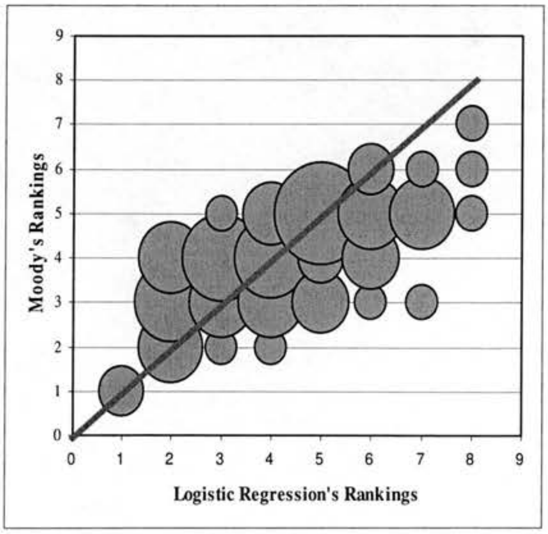


Figure 1. Logistic Regression vs. Moody's (2002)

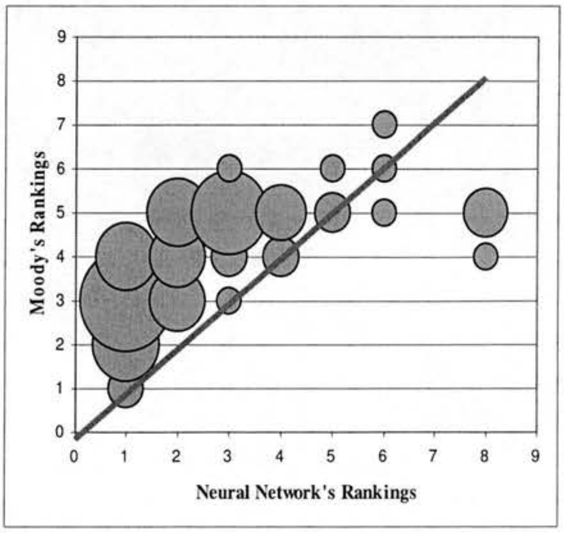


Figure 2. Neural Network vs. Moody's (2001)

The bubble chart analysis is a quick way of comparing the degree of similarity between different ranking methods. Comparing the two figures above, the better performing model, logistic regression model, has more bubbles with large size close to the diagonal line than the worse performing model, neural network. Also, the neural network vs. Moody's diagram has a greater number of large bubbles away from the diagonal line.

Overall, it can be concluded that logistic regression model's ability to rank company performance is slightly better than SMO model and is relatively better than the neural network model. This is true for significance testing using Kolmogorov-Smirnov test, distance analysis, and agreement measure using weighted Kappa. Moreover, the bubbles distribution is a very useful graphical method to detect the similarity between two ordinal datasets.

8. Discussions and Further Research

This paper proposed a theoretical framework for predicting financial distress based on Hunt's (2000) Resource-Advantage (R-A) Theory of Competition. 170 measures were drawn from literature on performance measurement and interviews with outside stakeholders. After a regrouping process, 67 variables are chosen out of the 170 for model construction and key variables were found via cluster analysis, univariate analysis and forward stepwise approach.

The USA retail sector was also chosen because of the clear definitions and reporting of financial distress through Chapters 7 and 11. Five credit scoring methodologies: Naïve Bayes, Logistic Regression, Recursive Partitioning, Artificial Neural Network, and Sequential Minimal Optimization, were used on a sample of 195 healthy companies and 51 distressed firms over five time periods from 1994 to 2002.

The time scale analysis showed unsurprisingly that all models with the time period one year prior financial distress show the best classification. Furthermore, even if the time period is five years prior to financial distress, the accuracy rates are above 80% and the AUROC values are above 0.80 among all five modelling methodologies. However, it is difficult to conclude that which modelling methodology has the absolute best classification ability, since the model's performance varies in terms of different time scales and different variable groups.

Regarding the external influences detection, this research showed that the external influences exist in all five credit scoring models, but these influences are weak. Furthermore, G12 models in the time period of one year before financial distress showed the best performance in terms of both accuracy rates (above 90%) and AUROC values (above 0.93).

The above results are potentially overly optimistic, due to the limitation of the sample. To overcome this problem, a series of comparison analysis from the study with Moody's rating were performed. Using the Kolmogorov-Smirnov significance test, distance measure, and weighted Kappa measure, it was found that logistic regression model's ability to rank company performance according to Moody's rankings is slightly better than SMO model and is relatively better than the neural network model. The bubbles distribution was also introduced in this research for detecting the similarity between two ordinal datasets and also presented similar results with other comparison techniques.

From the findings above, it can be argued that neural network model showed similar performance with logistic regression and SMO model based on the classification ability, but performed worse in terms of the comparison analysis with Moody's rating. An explanation is that neural network model fit closely to the sample and hence overfitting, whilst logistic regression and SMO models have not done so.

Finally, it must be noted that the scope of this study was limited to publicly listed firms and the US retail market. The study can be extended to non-listed firms as well as other markets in retail in order to ensure each model's theoretical utility and practical applicability.

References

Altman, E.I., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, Journal of Finance 23 (4) 589-609.

Altman, E.I., Haldeman, R.C., Narayanan, P., 1977. Zeta analysis: A new model to identify bankruptcy risk of corporations, Journal of Banking and Finance 1 (1) 29-54.

Beaver, W.H., 1966. Financial ratios as predictors of failure, Journal of Accounting Research 4 (3) 71-111.

Blum, M., 1974. Failing company discriminant analysis, Journal of Accounting Research 12 (1) 1-25.

Bowen, R.M., Daley, L.A., Huber Jr., C.C., 1982. Evidence on the existence and determinants of inter-industry differences in leverage, Financial Management 11 (4) 10-20.

Carson, J.M., Hoyt, R.E., 1995. Life insurer financial distress: classification models and empirical evidence, Journal of Risk and Insurance 62 (4) 764-775.

Casey, C., Bartczak, N., 1985. Using operating cash flow data to predict financial distress: some extensions, Journal of Accounting Research 23 (1) 384-401.

Coats, P.K., Fant, L.F., 1993. Recognizing financial distress patterns using a neural network tool, Financial Management 22 (3) 142-155.

Cohen, J. 1960. A coefficient of agreement for nominal scales, Educational and Psychological Measurement 20 37-46.

Dawson, J., 2000. Retailing at century end: some challenges for management and research, International Review of Retail, Distribution and Consumer Research 10 (2) 119-148.

Deakin, E.B., 1972. A discriminant analysis of predictors of business failure, Journal of Accounting Research 10 (1) 167-179.

Deakin, E.B., 1976. Distributions of financial accounting ratios: some empirical evidence, Accounting Review 51 (1) 90-96.

Dimitras, A.L., Slowinski, R., Sumaga, R., Zopounidis, C., 1999. Business failure prediction using rough sets, European Journal of Operational Research 114 (2) 263-280.

Edmister, R.O., 1972. Financial ratios as discriminant predictors of small business failure, Journal of Finance 27 (1) 139-140.

Eisenbeis, R.A., 1977. Pitfalls in the application of discriminant analysis in business, finance and economics, Journal of Finance 32 (3) 875-900.

Fan, A., Palaniswami, M., 2000. Selecting bankruptcy predictors using a support vector machine approach, Paper presented at 2000 IEEE-INNS-ENNS International Joint Conference on Neural Networks.

Fisher, R.A., 1936. The use of multiple measurements in taxonomic problems, Annals of Eugenics 7 179-188.

Fitch Ratings, 2000. Assigning Credit Ratings to European Retails, Fitch press, New York. (Downloadable from website http://www.fitchratings.com/)

Fitch Ratings, 2001. Corporate: Corporate Rating Methodology, Fitch press, New York. (Downloadable from website http://www.fitchratings.com/)

Frydman, H., Altman, E.I., Kao, D.L., 1985. Introducing recursive partitioning for financial classification: the case of financial distress, Journal of Finance 40 (1) 269-291.

Gentry, J.A., Newbold, P., Whitford, D.T., 1985. Classifying bankrupt firms with funds flow components, Journal of Accounting Research 23 (1) 146-160.

Gupta, M.C., Huefner, R.J., 1972. A cluster analysis study of financial ratios and industry characteristics, Journal of Accounting Research 10 (1) 77-95.

Hamer, M.M., 1983. Failure prediction: Sensitivity of classification accuracy to alternative statistical methods and variable sets, Journal of Accounting and Public Policy 2 189-307.

Hand, D.J., 1997. Construction and Assessment of Classification Rules, John Wiley & Sons Ltd, Chichester, England.

Hosmer, D.W., Lemeshow, S. 2000. Applied Logistic Regression, Wiley press, New York.

Hunt, S.D., 2000. A General Theory of Competition, Sage Publications Inc, California.

Hunt, S.D., Morgan, R.M., 1997. Resource-Advantage theory: a snake swallowing its tail or a general theory of competition?, Journal of Marketing 61 (4) 74-82.

Libby, R., 1975. Accounting ratios and the prediction of failure: some behavioral evidence, Journal of Accounting Research 13 (1) 150-161.

Mar-Molinero, C., Serrano-Cinca, C., 2001. Bank failure: a multidimensional scaling approach, European Journal of Finance 7 (2) 165-183.

Marais, M.L., Patell, J.M., Wolfson, M.A., 1984. The experimental design of classification models: An application of recursive partitioning and bootstrapping to commercial bank loan classifications, Journal of Accounting Research 22 (Supplement) 87-114.

McKee, T.E., Lensberg, T., 2002. Genetic programming and rough sets: A hybrid approach to bankruptcy classification, European Journal of Operational Research 138 (2) 436-451.

Mensah, Y.M., 1983. The differential bankruptcy predictive ability of specific price level adjustments: some empirical evidence, Accounting Review 58 (2) 228-246.

Mensah, Y.M., 1984. An examination of the stationarity of multivariate bankruptcy prediction models: A methodological study, Journal of Accounting Research 22 (1) 380-395.

Moody's Investors Service Inc., 1998. Rating Methodology: Industrial Company Rating Methodology, Moody's Investors Service Inc. Press, New York. (Downloadable from website http://www.moodys.com/)

Moody's Investors Service Inc., 2002. Retail Rating Methodology: Moody's Approach to Assessing Key Credit Issues in Retailing, Moody's Investors Service Inc. Press, New York. (Downloadable from website http://www.moodys.com/)

Moore, A.W., 2001. Cross-validation for detecting and preventing overfitting, Carnegie Mellon University (Downloadable from website: http://www.cs.cmu.edu/~awm/381/2004/crossvalidation.pdf)

Ohlson, J.A., 1980. Financial ratios and the probabilistic prediction of bankruptcy, Journal of Accounting Research 18 (1)109-131.

Platt, H.D., Platt, M.B., 1990. Development of a class of stable predictive variables, Journal of Business Finance & Accounting 17 (1) 31-51.

Platt, J.C., 1999. Fast training of support vector machines using sequential minimal optimization, In Schölkopf, B., Burges, C.J.C., Smola, A.J., 1999. Advances in Kernel Methods: Support Vector Machines, MIT press, Cambridge, England

Rose, P.S., Andrews, W.T., Giroux, G.A., 1982. Predicting business failure: A macroeconomic perspective, Journal of Accounting, Auditing & Finance 6 (1) 20-31.

Sanger, G.C., McConnell, J.J., 1986. Stock exchange listings, firm value, and security market efficiency: The impact of NASDAQ, The Journal of Financial and Quantitative Analysis, 21 (1) 1-25.

Standard and Poor's, 2002. Standard and Poor's 2002 Corporate Rating Criteria, The McGraw-Hill Companies press, New York. (Downloadable from website: http://www.standardandpoors.com/)

Standard and Poor's, 2003. Standard and Poor's 2003 Corporate Rating Criteria, The McGraw-Hill Companies press, New York. (Downloadable from website: http://www.standardandpoors.com/)

Taffler, R.J., 1984. Empirical models for the monitoring of UK corporations, Journal of Banking and Finance 8 (2)199-227.

Tam, K.Y., Kiang, M.Y., 1992. Managerial applications of neural networks: the case of bank failure predictions, Management Science 38 (7) 926-947.

Thomas, L.C., Edelman, D.B., Crook, J.N., 2002. Credit Scoring and Its Applications, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Trigueiros, D., Taffler, R. 1996. Neural networks and empirical research in accounting, Accounting & Business Research 26 (4) 347-355.

Williams, W.H., Goodman, M.L., 1971. A statistical grouping of corporations by their financial characteristics, Journal of Financial and Quantitative Analysis.6 (4) 1095-1104.

Zhang, G.P., Hu, M.Y., Patuwo, B.E., Indro, D.C. 1999. Artificial neural networks in bankruptcy prediction: general framework and cross-validation analysis, European Journal of Operational Research 116 (1) 16-32.

Appendix A: Performance Measures Arrangement

Internal Resources Group

Resources	Principle	Main Measures
Financial Resources	Profitability	1. EBIT margin
		2. EBITDA margin
		3. EBITDAR margin
		4. Pre-tax profit margin
		5. Pre-tax profit on capital
		6. Net profit margin
		7. Gross profit margin
		8. SG&A as % of net sales
		9. EBIT on capital
		10. Return on total assets
		11. Return on total equity
		12. Operating margin
		13. Dividend payout ratio
	Liquidity	14. Current ratio
		15. Acid ratio
		16. Cash ratio
	Sustainability	17. Net operating cash flow / gross capex
		18. Cash dividend cover
		19. Fixed charge cover
		20. Interest cover
		21. Funds from operations / total debt
		22. EBITDA / interest
		23. Total debt / discretionary cash flow
	Leverage	24. Gearing ratio
		25. Debt / EBITDA
		26. Leased-adjusted net debt / EBITDAR
		27. Net debt / market capitalization
		28. Total debt / (total debt + market capitalization)
		29. Debt to equity ratio

Internal Resources Group (Continue)

Resources	Principle	Main Measures
Financial Resources	Market Measure	30. P/E ratio
	Financial Scale	31. Net sales 32. Total assets 33. Market share by retail sector (based on sales) 34. Market share by retail sector (based on gross margin) 35. Total capital employed 36. Operation cash flow
Physical Resources	Reach Ability	37. Store numbers
Legal Resources	Brand Strength	38. Market capitalization / net assets
Human Resources	Human Resource Quality	39. Sales per employee 40. EBIT per employee
	Human Resource Management	41. Number of payrolls
Organizational Resources	Actability	42. Total assets turnover 43. Fixed assets turnover
	Growth Power Analysis	44. Sales growth 45. Market value growth 46. Capital growth 47. EBIT growth 48. Number of stores growth 49. The operating income growth 50. Number of payrolls growth
	Financial Management	51. Net cash cycle
Informational Resources	Market Segment Risk management	52. Main market sales as percentage total sales
Relational Resources	Customer Relations Management	53. Receivable turnover
	Supplier Relations Management	54. Inventory turnover 55. Payables turnover

External Environmental Factors

Factors	Principle	Main Measures
Societal Resources Societal Institutions Government Actions	Political Environmental Factors	56. The correlation coefficient between government debt / GDP and total sales 57. The correlation coefficient between government avenue / GDP and total sales 58. The correlation coefficient between government expense / GDP and total sales
	Economic Environmental Factors	59. The correlation coefficient between real GDP and total sales 60. The correlation coefficient between average interest rate and total sales 61. The correlation coefficient between unemployment rate and total sales 62. The correlation coefficient between disposable income and total sales
	Technological Environmental Factors	63. The correlation coefficient between total government spending for R&D and total sales
	Socio-cultural Environmental Factors	64. The correlation coefficient between birth rate and total sales 65. The correlation coefficient between death rate and total sales 66. The correlation coefficient between age structure ratio (0-14 years old) and total sales 67. The correlation coefficient between age structure ratio (65 years and above) and total sales

Paper Two:

Developing Financial Distress Prediction Models: A Study of US, Europe and Japan Retail Performance

- Presented at *Taiwan Scientific Symposium in Edinburgh*, ETSS and Science and Technology Division, Taipei Representative Office in the UK, University of Edinburgh, Feb. 18 2006
- Presented in the Proceedings of *The 4th Edinburgh Conference on RISK: Analysis, Assessment and Management*, The Edinburgh University Management School and Economics, Business and Industrial Section and Royal Statistical Society, Mar. 29 – Mar. 31 2006.
- Presented in the Proceedings of *2006 Financial Management Association (FMA) European Conference*, Stockholm, Sweden, June 7 – June 10
- Presented in the Proceedings of *2006 European Financial Management Association (EFMA) Annual Conference*, Madrid, Spain, June 28 – July 1
- Presented in the Proceedings of *2006 International Conference in Business and Information (BAI)*, Singapore, July 12 – July 14
- Presented in the Proceedings of *The joint 14th annual Pacific Basin Financial Economics and Accounting and 2006 annual Financial Engineering Association of Taiwan Conference*, Taiwan, July 14 – July 15

Developing Financial Distress Prediction Models

A Study of US, Europe and Japan Retail Performance

Yu-Chiang Hu[a,1] and Jake Ansell[a]

[a]*Management School and Economics, University of Edinburgh, William Robertson Building, 50 George Square, Edinburgh, EH8 9JY, UK*

Abstract

This paper constructs retail financial distress prediction models based on five key variables previously shown to have good classification properties (Hu and Ansell, 2005). Five credit scoring techniques—Naïve Bayes, Logistic Regression, Recursive Partitioning, Artificial Neural Network, and Sequential Minimal Optimization (SMO) were considered. A sample of 491 healthy firms and 68 distressed retail firms were studied over a five-year time period from 2000 to 2004.

An international comparison analysis of three retail market models –USA, Europe and Japan– shows that the average accuracy rates are above 86.5% and the average AUROC values are above 0.79. Almost all market models display the best discriminating ability one year prior to financial distress. The US market model performs relatively better than European and Japanese models five years before financial distress.

A composite model is constructed by combining data from US, European and Japanese markets. All five credit-scoring techniques have the best classification ability in the year prior to the financial distress, with accuracy rates of above 88% and AUROC values of above 0.84. Furthermore, these techniques still remain sound five years before financial distress, as the accuracy rate is above 85% and AUROC value is above 0.72. However, it is difficult to conclude which modelling technique has the absolute best classification ability, since the composite model's performance varies according to different time scales.

Regarding the applicability of the composite model, a comparison is made using Moody's credit ratings. Results indicate that SMO is the better performing model amongst the three models, closely followed by the neural network model. Logistic regression model shows lowest performance in terms of similarity with Moody's.

Keywords: Credit Risk, Financial Distress Prediction; Multivariate statistics; Artificial intelligence

[1] Corresponding author

Email address: Y.A.Hu@sms.ed.ac.uk (Yu-Chiang Hu), Jake.Ansell@ed.ac.uk (Jake Ansell)

* Acknowledgement: This research is funded by the College of Humanities and Social Science, Management School and Economics at University of Edinburgh as well as the Overseas Research Students Awards Scheme (ORSAS).

1. Introduction

How can financial distress be predicted? This question is of interest not only to managers but also to external stakeholders of a company. These players are continuously seeking the optimal solution for performance forecasting, as a way to rationalize the decision-making process. Thus, the primary objective of this paper is to establish financial distress prediction models based on credit-scoring techniques.

A single industry is chosen to avoid generalizations across industries. The retail industry is selected, as assessing and evaluating retail risk is one of the key issues in retail research (Dawson, 2000). Variable selection is derived from findings in Hu and Ansell (2005). Based on a USA retail dataset of 195 healthy firms and 51 distressed firms for years 1994 to 2002, Hu and Ansell (2005) showed that five critical performance variables: *Debt Ratio*, *Total Debt / (Total Debt + Market Capitalization)*, *Total Assets*, *Operating Cash Flow* and *Government Debt / GDP* have sound classification ability (accuracy rate of above 90% and AUROC value of above 0.935) one year before financial distress. Moreover, even if the time period is five years prior to financial distress, the classification accuracy rate using these variables is above 80% and the AUROC value is above 0.80.

This research employs five credit-scoring techniques: *Naïve Bayes*, *Logistic Regression*, *Recursive Partitioning*, *Artificial Neural Network*, and *Sequential Minimal Optimization (SMO)* for modelling purposes. Three target markets, USA, Europe and Japan, are selected for an international comparison analysis. Comparative results show that regardless of the target countries, the average accuracy rates are above 86.5% and the average AUROC values are above 0.79. Moreover, exploring the time dimension, all three market models perform best in the year prior to financial distress with slight difference across markets. However, the longer the period before financial distress, the greater the difference across markets becomes, especially in terms of the AUROC values. For example, five years prior to financial distress, the US has significantly better AUROC value than Japan or Europe.

The research develops a composite model based on a sample of 491 healthy and 68 distressed retail firms over the time period from 2000 to 2004 by combining data from the USA, Europe and Japan. Results show that all five credit-scoring techniques in the year prior to the financial distress display the best performance with accuracy rates of above 88% and AUROC values of above 0.84. Furthermore, these techniques still remain sound five years before financial distress, as the accuracy rate is above 85% and AUROC value is above 0.72.

However, it is difficult to conclude which modelling methodology has the absolute best classification ability, since the model's performance varies in terms of different time scales.

Finally, in order to examine potential overfitting problems in the composite model, a comparison of the composite model with Moody's credit rating is carried out. The results indicate that SMO is the better performing model amongst the three models, closely followed by neural network model. Logistic regression model shows lowest performance in terms of similarity with Moody's.

Financial distress prediction modelling techniques will be discussed in section 2. Section 3 will illustrate the variable selection and data collection. Section 4 describes the methodologies employed to evaluate modelling utility and compare results with Moody's rating. The results will be analyzed in section 5. Finally, a discussion of the results will be presented in section 6.

2. The Development of Default Prediction Modelling Techniques

Financial distress prediction became a critical accounting and finance research area since 1960s. Based on the cash flow framework, Beaver carried out three different univariate analyses—profile analysis (comparison of mean values), dichotomous classification test and likelihood ratio analysis— in order to examine the predictive characterises and utility of each variable. Regarding the likelihood ratio analysis, Beaver (1966) conducted an analysis of likelihood ratios based on the *Bayesian* approach. He argued that the default prediction problem could be regarded as a problem of evaluating the probability of financial distress conditional upon the value of a specific financial ratio. He further pointed out that financial ratios can provide useful information for predicting default, since the likelihood ratios still present high values even five years prior to financial distress. Let D represents the distressed sample and X is the vector of independent variables and assume x is a particular vector of an independent variable. The conditional probability of a financial distress company in terms of a specific financial ratio x can be expressed as:

$$P(D \mid X = x) = \frac{P(D)P(X = x \mid D)}{P(X = x)} \tag{1}$$

Univariate analysis is limited in the evaluation of a firm's performance, since it is difficult to use only one single measure to describe the performance in a multidimensional firm.

However, prior to construct a multivariate model, it is still useful to carry out a univariate analysis for the purpose of variable selection, as not every variable has good discriminating utility (Hosmer and Lemeshow, 2000).

Altman (1968) was the first researcher to apply the *Multiple Discriminant Analysis (MDA)* approach to the financial distress prediction domain. He developed a Z-score bankruptcy prediction model and determined a cutpoint of Z-score (2.675) to classify healthy and distressed firms. The results showed that the Z-score model had sound prediction performance one year and two years before financial distress, but did not indicate good prediction utility three to five years before financial distress. A number of authors followed his work, and applied the Z-score model into different markets, different time periods and different industries, such as, Taffler (1982, 1984), Pantalone and Platt (1987), Betts and Belhoul (1987) and Piesse and Wood (1992).

However, MDA assumes that the covariance matrices of two populations are identical and both populations need to be described by multivariate normal distribution. Clearly, these assumptions do not always reflect the real world. Deakin (1976) argued that even if after performing the normality transforming process, financial ratio data do not follow normal distribution. Moreover, Hamer (1983) evaluated the sensitivity of financial distress prediction models in terms of four different variable sets from previous research (Altman, 1968; Deakin, 1972; Blum, 1974; Ohlson, 1980) and she pointed out that the covariance matrices in each variable set were statistically different.

Ohlson (1980) was the first to apply the *Logistic Regression* model to financial distress prediction research. After Ohlson's (1980) work, the conditional probability model became a popular modelling technique in the bankruptcy prediction domain (also see Zavgren, 1983; Mensah, 1983; Casey and Bartczak, 1985) The logistic regression model can be linearized by logit transformation on odd ratio function and can be expressed as follow:

$$g(x) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n \qquad (2)$$
$$= \beta \times x^T$$

Where $\pi(x)$ is the logistic function,

$$\pi(x) = \frac{1}{1+e^{-(\beta \times x^T)}} = \frac{e^{\beta \times x^T}}{1+e^{\beta \times x^T}} \qquad (3)$$

Although logistic regression does not suffer from the limitations of MDA, Tabachnick and Fidell (2000) pointed out that if the assumptions regarding the identical covariance matrices and multivariate normal distribution are met, MDA is likely to be more efficient than logistic regression. Moreover, like all the regression functions, the problem of multicollinearity still exists in logistic regression.

Recursive Partitioning (RP) was introduced in the bankruptcy prediction research in the mid-1980s (Marais et al., 1984; Frydman et al., 1985). RP is a non-parametric technique and does not suffer the limitations from traditional statistical models. Based on the lowest expected misclassification cost, RP first selects an independent variable as the best discriminator and decides a cutpoint. The next step is to classify both healthy and distressed firm into two sub-nodes in terms of the cutpoint. The third step is to select another (or the same) discriminator and further partition the healthy and distressed firms into another two sub-nodes. The same process can be continued, if further splitting is necessary. It is obvious that the overfitting may be a potential problem of RP, since the continuous partitioning process is likely to encourage one misclassified case in the terminal node. Therefore, Thomas et al. (2002) pointed out that if the sample size in a node is too small, then further partition is not appropriate. Moreover, if the classification difference between the old node and new nodes is not significant, the partitioning process is not necessary to continue.

From the late 1980s, the Machine Learning techniques in the Artificial Intelligence (AI) area, such as *Artificial Neural Networks (ANN)*, were applied to financial distress prediction studies (Coates and Fant, 1993; Zhang et al., 1999). The most popular ANN algorithm in the financial distress prediction domain is the *Multilayer Perceptron (MLP)*. The composition of MLP has three main components: input layer, hidden layer and output layer, and they are illustrated in the Figure 1.

The ANN training process can be regarded as a weighting determination process. The most frequently used algorithm for training process is the *Back Propagation Algorithm (BPA)*. Thomas et al. (2002) mentioned that BPA first calculates the difference between the expected output value and the observed output value (called *error*) in the output layer and then distributes the error back to the network with a weight. The next step is to adjust the weight to reduce the error. The same process is repeated for all cases, called an *epoch*. After several epochs training, the learning error will reduce to a minimum level and the training process ends. Trigueiros and Taffler (1996) mentioned some advantages of MLP. For example, as recursive partitioning, it does not require the statistical distribution assumptions. However, MLP still has some limitations, such as no adequate significance tests and requirement of computer power (Tam and Kiang, 1992).

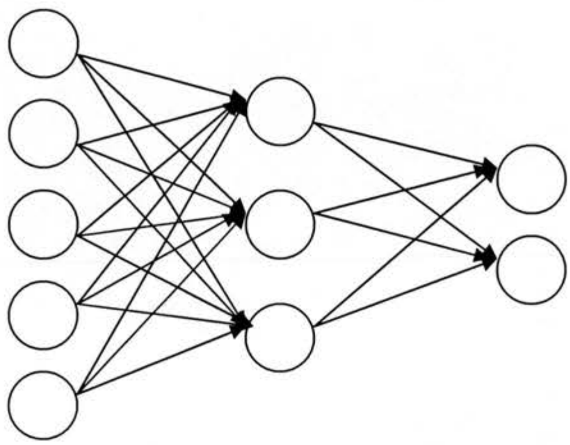


Figure 1 Three Layers Multilayer Perceptron

The input layer is responsible for receiving information from the outside environment and transferring it to the hidden layer. In the hidden layer, a neuron will assign a series of weights to the inputs, cope with the information via a training process, and then forward the results with weights to the output layer.

In the late 1990s, *Support Vector Machine (SVM)* was introduced to cope with the classification problem. Fan and Palaniswami (2000) applied SVM to select the financial distress predictors. They pointed out that SVM created an optimal separating hyperplane in the hidden feature space in terms of the principle of structure risk minimization and used the quadratic programming to obtain an optimal solution. However, Platt (1999) argued that a large number of quadratic programming in SVM training is time consuming. As a result, he introduced a new algorithm, *Sequential Minimal Optimization (SMO)*, to improve the SVM training time, since SMO only uses two Lagrange multipliers at each training step. Plat (1999) also pointed out that SMO has better performance than other SVM training methods in terms of many aspects, such as better scaling with training sample size. From the early 2000s, some other credit scoring modelling techniques were also employed in the bankruptcy prediction research area and have shown good performance, including the *Rough Sets* approach (McKee, 2003) and the *Multidimensional Scaling* approach (Mar-Molinero and Serrano-Cinca, 2001).

3. Variable Selection and Data Collection

3.1 Variable Selection

Hu and Ansell (2005) developed a theoretical framework for retail performance measure selection based on Hunt's (2000) Resource-Advantage (R-A) Theory of Competition and 170 potential retail performance measures, which cover both internal and external measures, were obtained in terms of previous literature survey and interviews with outside stakeholders. In this framework, Hu and Ansell (2005) considered several important aspects relative to variable selection in the previous bankruptcy prediction studies. For example, some studies pointed out that the macro-economical factors have great impacts on a default prediction model (Rose et al., 1982; Mensah, 1984). Hu and Ansell (2005) considered the external variables not only based on the economical environment, but also took into account the political environment, social-culture environment and technological environment.

Moreover, they also considered the qualitative performance measures in terms of the practical point of view, since many renowned credit-rating companies including Moody's, S&P, and Fitch consider both quantitative and qualitative factors when carrying out credit evaluation but attribute more importance to qualitative rather than quantitative factors in the process (Moody's, 1998 and 2002; Fitch, 2000 and 2001; S&P, 2002 and 2003). In addition, the industrial variables, such as store number, were also contemplated in their framework, since some authors argued that the industry-relative measures could improve the accuracy of the classification model (Platt and Platt, 1990).

Although Hu and Ansell (2005) took into account many potential performance measures, it is obvious that too many variables in a prediction model tend to overfit the model utility, and hence provides a subjective conclusion. Drawing on this insight, they selected key performance measures by using the logistic forward stepwise analysis. In addition, prior to select the final variables, some key issues include: time-scale consideration, outlier elimination and univariate analysis, were carried out in order to ensure the quality of key variables. The results provided sufficient evidence that that these five variables have sound classification ability (accuracy rate is above 90% and AUROC value is above 0.935) one year prior to financial distress. Furthermore, even if the time period is five years prior to financial distress, the classification accuracy rate using these variables is above 80% and the AUROC value is above 0.80. These five key variables are illustrated as follows:

(1) Leverage Measures: Debt Ratio and Total Debt / (Total Debt + Market Capitalization)

Debt Ratio and Total Debt / (Total Debt + Market Capitalization) are used to evaluate a company's leverage situation, especially to measure a company's ability to face its long-term obligations. Therefore, these two measures are related to a company's credit assessment directly. One of the differences between these two measures is equity evaluation. For debt ratio (total debts / total assets), the value of equity is evaluated by accounting value, whilst the equity value is evaluated by market value for another leverage measure.

Another difference between these two measures is the maximum value. For the ratio of total debt / (total debt + market capitalization), the maximum value is one, since the minimum value of the market capitalization is zero. However, for the debt ratio, the maximum value is possible to greater than one, since the value of total debt is possible to greater than the value of total assets. It implies that even if a company sell all its assets, this company still cannot cover their future obligations. In fact, if a company's debt ratio is greater than one, this company is under the stock-based insolvency situation (Altman 1983, Ross, et al. 1999). In other words, this company is currently facing financial distress.

However, a higher leverage may not mean a higher bad debt risk, since it depends on the composition of the leverage. Fitch (2000) argued that distinguishing the financial leverage and operating leverage is very important in the retail industry, since the operating leverage, such as the loan for store equipments purchasing, is caused by the customer's demand, and hence not so risky. Drawing on this insight, leverage analysis should focus on financial leverage rather than operating leverage.

(2) Scale Measure: Total Assets

Scale measures are more important in the retail industry than in other industries, as one of the important characteristics in the retail industry is low-margin. Large firms usually have certain advantages, which small firms do not have. For example, large firms have better risk endurance when the economical situation changes. Moreover, large firms also have better financial flexibility, since they can more easily ask for a loan from a financial institution than small firms (S&P, 2003). As a result, size is a significant variable for evaluating a retailer's credit risk.

(3) Sustainability Measure: Operating Cash Flow

Sustainability measures a company's ability to service external sources of finance, such as interest payments. S&P (2003) pointed out that a company's sustainability must be based on cash flow, rather than on earnings in the accounting statements, for earnings include non-cash items that cannot reflect a company's ability to pay back future obligations. Thus, if a company has adequacy operating cash flow, the default risk will be lower.

(4) External Environmental Measure: Government Debt / GDP

Government debt / GDP can be regarded as a measure to evaluate a country's leverage situation, since it indicates the ability of a country to cover its total debt by using GDP. Therefore, this measure is usually applied to evaluate a country's sovereign risk (S&P, 2005). In order to assess this measure's impacts on each sample company, a five years correlation coefficient between government debt / GDP and total sales is employed.

3.2 Sample Selection Criteria

Regarding the sample selection of healthy firms, two criteria are considered. Only listed firms are considered, since listed companies need to obey the regulations in the financial market, their data are more transparent. Another important sample selection criterion is that this research does not consider e-retailers, because the performance measures of e-retailers are different. Finally, even if a company satisfied the criteria above, it was excluded if its data is not complete.

In connection with the sample selection of distressed companies, the criteria are based on the financial point of view. Ross, et al. (1999) pointed out the definition of financial distress has two themes: stock-based insolvency and flow-based insolvency. Stock-based insolvency occurs when a company's total liabilities are greater than its total assets. Flow-based insolvency occurs when a company's operating cash flow cannot meet its routine obligations. Hence, a company was regarded as distressed in this research if its debt to equity ratio was negative or if its interest cover based on cash flow framework (EBITDA / interest) was smaller than one.

3.3 Data Collection

Thomson One Banker database was the main data source of each company's financial data. The macroeconomical data was collected from the documents in the Organisation for Economic Co-operation and Development (OECD). Table 1 summarised the data collection results in terms of three target markets: USA, Europe and Japan, from 2000 to 2004.

Table 1 Overall Data Description

	2004		2003		2002		2001		2000	
	Healthy	Distressed	Healthy	Distressed	Healthy	Distressed	Healthy	Distressed	Healthy	Distressed
USA	181	24	179	40	190	46	184	63	190	70
European[2]	145	27	162	26	164	31	182	32	195	31
Japan	251	28	244	19	219	17	180	55	195	39
Total	577	79	585	85	573	94	546	150	580	140

An initial interest of this study was the timescale effect, whether one should use data just before the default or some time before. Hence, this research adapted 2004 as the year prior to financial distress, and then it allowed series timescale effect detection. For example, 2003 can be regarded as the time period two years before financial distress; 2002 can be viewed as the time period three years before financial distress, and so on. As a result, only the sample company, which has five years complete data, was considered for exploring timescale effect. The sample size of each country is illustrated in Table 2:

Table 2 Data description for exploring time scale effect

	USA	European	Japan	Total
Healthy	170	126	195	491
Distressed	21	20	27	68

4. Methodology

Prior to model construction, a cross-validation process was performed to solve overfitting problem and the 10-folders approach was selected for the purpose of cross-validation. Moore (2001) compared three cross-validation methods: the test set method, the leave one out method and the 10-folders method and argued that the 10-folders cross-validation approach only wasted 10% of total data and the training cost was lower than the leave one out method.

2 The composition of the European market includes the 25 countries in the European Union plus Swaziland and Norway.

Five credit scoring techniques are employed for model construction: *Naïve Bayes*, *Logistic Regression*, *Recursive Partitioning*, *Artificial Neural Network* and *Sequential Minimal Optimization (SMO)*. Model classification ability was evaluated in terms of two approaches: the *Classification Accuracy Rate* approach and the *Area under the Receiver Operating Characteristics Curve (AUROC)* approach. Classification accuracy rate is a straightforward method employed widely in previous studies on model evaluation. The area under the ROC curve (AUROC) is the area between the ROC curve and the diagonal line and hence the value of AUROC is between 0.5 and 1. The diagonal line of ROC curve reflects the feature of a test with no discriminating power, (Hand, 1997). In fact, different cut points should reflect different sensitivity and specificity values, since the classification rule is different. Therefore, the further the ROC curve is from the diagonal line, the better the model performance (Thomas et al., 2002). In this research, AUROC is applied to the naïve bayes, logistic regression and artificial neural network models.

Given the sample size available for study it was not possible, and probably it would not have been informative, to employ a hold out sample. Hence, the above methodology will result in potentially overly optimistic results. To overcome this problem for the best modelling approaches, it was decided to compare the credit scores from the composite model with a standard rating system; in this case Moody's rating. In retailing, there are only 8 rating grades given Aa to C in Moody's system. Hence, the data was ranked according to score and divided into 8 groups. Logistic regression, neural network and SMO models are selected for the ranking comparison analysis. A range of measures for comparison were used, *Kolmogorov-Smirnov (K-S) Test*, *Distance Analysis*, and *Weighted Kappa Analysis* and finally *Graphical Bubble Charts*.

5. Empirical Analysis

5.1 International Comparison Analysis

Based on the data in Table 1, an international comparison analysis in terms of both the accuracy rate and AUROC value can be carried out. Table 3 presents the results in different countries over 5 years period. It is very obvious that regardless of the target countries, the average accuracy rates are above 86.5% and the average AUROC values are above 0.79. The results suggest that the five key variables have sound prediction ability in American, European and Asian retail markets.

Table 3 Model Performance in Target Markets

USA Market							
Methodology	**Measures**	**2004**	**2003**	**2002**	**2001**	**2000**	**Average**
Naïve Bayes	Accuracy Rate (%)	89.76	90.41	91.10	87.85	86.54	***89.13***
	Average AUROC	0.9332	0.9345	0.9418	0.9180	0.9238	***0.9303***
Logistic Model	Accuracy Rate (%)	92.20	90.87	90.25	87.85	87.69	***89.77***
	Average AUROC	0.9399	0.9426	0.9141	0.9137	0.9168	***0.9254***
Neural Network	Accuracy Rate (%)	91.22	87.21	88.14	87.04	85.77	***87.88***
	Average AUROC	0.8946	0.8997	0.8715	0.8992	0.8944	***0.8919***
SMO	Accuracy Rate (%)	92.68	89.95	89.83	87.04	82.31	***88.36***
Recursive Partitioning	Accuracy Rate (%)	93.17	88.58	88.56	82.59	83.85	***87.35***
European Market							
Methodology	**Measures**	**2004**	**2003**	**2002**	**2001**	**2000**	**Average**
Naïve Bayes	Accuracy Rate (%)	88.95	89.89	84.62	87.85	89.38	***88.14***
	Average AUROC	0.8835	0.8184	0.8387	0.9250	0.8996	***0.8730***
Logistic Model	Accuracy Rate (%)	90.12	88.30	90.26	90.19	90.71	***89.92***
	Average AUROC	0.8733	0.8253	0.8324	0.9050	0.8951	***0.8662***
Neural Network	Accuracy Rate (%)	88.95	87.77	83.59	90.65	92.04	***88.60***
	Average AUROC	0.8248	0.8029	0.7710	0.8913	0.9179	***0.8416***
SMO	Accuracy Rate (%)	88.95	86.70	84.62	85.51	88.05	***86.77***
Recursive Partitioning	Accuracy Rate (%)	91.28	89.36	85.13	87.38	87.61	***88.15***
Japan Market							
Methodology	**Measures**	**2004**	**2003**	**2002**	**2001**	**2000**	**Average**
Naïve Bayes	Accuracy Rate (%)	89.25	91.25	91.10	79.57	84.62	***87.16***
	Average AUROC	0.8552	0.8013	0.8020	0.8545	0.8368	***0.8300***
Logistic Model	Accuracy Rate (%)	90.32	91.64	94.49	80.00	87.18	***88.73***
	Average AUROC	0.8314	0.8218	0.7814	0.8515	0.8393	***0.8251***
Neural Network	Accuracy Rate (%)	89.96	92.02	93.22	78.30	88.03	***88.31***
	Average AUROC	0.8059	0.7677	0.7806	0.8098	0.7888	***0.7906***
SMO	Accuracy Rate (%)	89.96	92.78	92.80	78.72	88.46	***88.54***
Recursive Partitioning	Accuracy Rate (%)	89.61	93.16	92.37	76.60	88.03	***87.95***

In addition, the naïve bayes model and SMO model show the best performance in the US market, whilst the recursive partitioning model displays the best performance in the European market. The logistic regression model and the neural network model for the US market shows the best performance in terms of the average AUROC value, whereas The European market displays the best performance based on the average accuracy rate. Although the results show that the model performance is different in each country, the difference is very small. (The only exception is the performance of neural network model between US and Japanese markets in terms of the average AUROC value: the difference is around 0.1). Hence, there is little difference in the models performance.

5.2 Exploring Time Scale

As mentioned in the Section 3, a five years time period was explored. Table 4 presents the results in different markets.

Table 4 Exploring Timescale in Target Markets

USA Market						
Methodology	**Performance Measures**	**2004**	**2003**	**2002**	**2001**	**2000**
Naïve Bayes	Accuracy Rate (%)	90.58	90.58	91.62	89.53	87.96
	Average AUROC	0.9238	0.8964	0.8454	0.7751	0.8210
Logistic Model	Accuracy Rate (%)	92.15	90.58	91.62	92.67	92.67
	Average AUROC	0.9241	0.921	0.8555	0.8123	0.8709
Neural Network	Accuracy Rate (%)	90.05	87.43	91.10	92.15	92.15
	Average AUROC	0.9087	0.8524	0.7714	0.7218	0.8339
SMO	Accuracy Rate (%)	91.62	91.10	91.10	92.67	91.62
Recursive Partitioning	Accuracy Rate (%)	92.75	92.67	90.05	90.58	91.62
Europe Market						
Methodology	**Performance Measures**	**2004**	**2003**	**2002**	**2001**	**2000**
Naïve Bayes	Accuracy Rate (%)	89.73	88.36	85.62	86.99	85.62
	Average AUROC	0.8964	0.7813	0.7369	0.6603	0.6067
Logistic Model	Accuracy Rate (%)	88.36	88.36	86.30	86.30	84.93
	Average AUROC	0.8694	0.6619	0.6571	0.6052	0.5619
Neural Network	Accuracy Rate (%)	91.10	88.36	85.62	87.67	82.88
	Average AUROC	0.8294	0.7298	0.6754	0.6159	0.5001
SMO	Accuracy Rate (%)	91.10	86.30	86.30	86.30	86.30
Recursive Partitioning	Accuracy Rate (%)	89.73	87.67	89.04	87.67	86.30
Japan Market						
Methodology	**Performance Measures**	**2004**	**2003**	**2002**	**2001**	**2000**
Naïve Bayes	Accuracy Rate (%)	88.74	86.49	86.49	84.68	85.14
	Average AUROC	0.8454	0.7909	0.7837	0.7417	0.6948
Logistic Model	Accuracy Rate (%)	87.84	86.04	88.29	86.94	86.94
	Average AUROC	0.8184	0.7649	0.7928	0.7358	0.7005
Neural Network	Accuracy Rate (%)	86.94	84.68	89.19	83.78	86.94
	Average AUROC	0.7725	0.6999	0.8342	0.6443	0.6615
SMO	Accuracy Rate (%)	87.84	87.84	87.84	87.84	87.84
Recursive Partitioning	Accuracy Rate (%)	85.14	84.68	87.84	87.84	86.94

The results provide sufficient evidence that for almost all modelling approaches, the model shows the best performance in the year before financial distress for the target markets. When comparing US, Europe and Japan market results for *each* credit scoring approach in 2004, the differential of results across markets is small. Nevertheless, it is interesting to note that the US results show the best classification ability for all credit scoring techniques based on both accuracy rate and AUROC value, except for neural network based on the accuracy rate.

Notwithstanding the small differential across markets the year before financial distress (2004), it should be said that the longer the period before financial distress, the greater the difference across markets becomes, especially in terms of AUROC values. For example, five years prior to financial distress (2000), the US has significantly better AUROC value than Japan or Europe.

5.3 Composite Model Performance Analysis

The primary objective of this research is to develop retail financial distress prediction models by using credit scoring techniques. Thus, this research constructed a composite model based on the sample size of 491 healthy firms and 68 distressed firms (see Table 2) by combining the data from USA, European and Japanese markets. The results of the composite model performance are illustrated in Table 5:

Table 5 Composite Model Performance

Composite model							
Methodology	**Measures**	**2004**	**2003**	**2002**	**2001**	**2000**	**Average**
Naïve Bayes	Accuracy Rate (%)	90.34	88.91	88.01	85.15	85.15	***87.51***
	Average AUROC	0.8781	0.8400	0.7972	0.7649	0.7202	***0.8001***
Logistic Model	Accuracy Rate (%)	91.23	88.55	88.55	88.91	87.84	***89.02***
	Average AUROC	0.8769	0.8300	0.7862	0.7538	0.7203	***0.7934***
Neural Network	Accuracy Rate (%)	89.09	88.01	86.94	89.80	86.40	***88.05***
	Average AUROC	0.8472	0.8017	0.7451	0.7363	0.7228	***0.7706***
SMO	Accuracy Rate (%)	88.19	88.19	88.01	87.84	87.84	***88.01***
Recursive Partitioning	Accuracy Rate (%)	89.98	86.58	88.37	88.37	87.48	***88.16***

Table 5 shows the same conclusion as previous time series analysis that all five credit scoring techniques have the best classification ability in the year prior to the financial distress, with accuracy rates of above 88% and AUROC values of above 0.84. Furthermore, these techniques still remain sound five years before financial distress, as the accuracy rate is above 85% and AUROC value is above 0.72.

With regards to performance of the modelling techniques, the conclusion is the same as Hu and Ansell (2005) that no modelling methodology has the absolute best classification ability, since the model's performance varies in terms of different time scales. For example, logistic regression model shows the best performance in 2004, but the same cannot be concluded in different time scales. Furthermore, if we focus on the average performance of each modelling technique, it is obvious that the performance among five credit scoring approaches is very similar. (The maximum difference of the average accuracy rate is only 1.5% and the maximum difference of the AUROC value is only 0.03)

Thus far, the findings above prove that the model has sound discriminating ability, even if the time period is five years before financial distress. However, due to the sample size limits, a holdout sample is not likely to employ in this research, and hence, the current results are potentially overly optimistic. In order to overcome this problem, logistic regression, neural network and SMO models in the year prior to financial distress are selected for the objective of credit score ranking comparison with Moody's credit rating results.

5.4 Test of Significance

The Kolmogorov-Smirnov test assesses whether two datasets differ significantly. A *p*-value is greater than 0.05 implies two samples come from a similar distribution. Results of the Kolmogorov-Smirnov Test are shown in Table 6.

Table 6 Two-sample Kolmogorov-Smirnov (K-S) test

Modelling Methodology	K-S	2004	2003	2002	2001	2000
Logistic Model	Z Value	1.620	1.852	3.163	4.938	5.401
	p-value	0.010	0.002	0	0	0
Neural Network	Z Value	3.858	3.626	3.009	2.315	1.620
	p-value	0	0	0	0	0.010
SMO	Z Value	1.080	1.620	2.237	2.237	2.469
	p-value	0.194	0.010	0	0	0

The highlighted *p*-values in Table 6 is non significant and indicate when a proposed model provides rankings similar to Moody's. The result shows that only SMO has similar rankings to Moody's in 2004. Significance testing is useful for determining whether or not there is similarity in ranking. The following techniques attempt to assess the level or degree of similarity.

5.5 Distance Analysis

The simplest approach to compare the degree of similarity between two ordinal data sets is distance analysis. The basic rule for distance analysis is that the smaller the distance between the rankings from Moody's and the composite model, the better the similarity of the composite model. To calculate distances, each cell in a crosstabulation table is presented as a proportion of the total sample size. The cell value is then multiplied by the value in the distance matrix and then, the resulting values are summed up. This gives an overall distance between Moody's model and each of the proposed composite models. Results are illustrated in Table 7.

Table 7 Overall Distances Results

Modelling Methodology	**2004**	**2003**	**2002**	**2001**	**2000**	***Average Distance***
Logistic Model	1.5595	1.7381	1.8095	2.5714	2.9167	***2.11904***
Neural Network	1.8810	1.9286	1.4286	1.2976	1.1786	***1.54288***
SMO	1.3929	1.3929	1.4762	1.5357	1.7381	***1.50716***

From Table 7, the logistic regression model shows the worst similarity measure, since its average distance is highest amongst the three models. In contrast, the SMO model displays the best performance. However, although neural network model's average distance is higher than SMO model, the difference is very small. Therefore, it can be concluded that the similarity of the SMO model is slightly better than the neural network model and relatively better than the logistic regression model.

5.6 Measure of Agreement

When companies are evaluated by different raters, it is important to measure the degree of agreement between these raters. The main question is that how much do the ratings provided by the logistic regression model, the neural network model, and the SMO model concord with those from Moody's? Weighted Kappa is useful to answer this question and it is an extension of Cohen's Kappa (1960) suitable for ordinal data and for measuring relative concordance. The values of weighted Kappa are presented in Table 8.

Table 8 Weighted Kappa Analysis (Previous)

Modelling Methodology	**2004**	**2003**	**2002**	**2001**	**2000**	***Average Weighted Kappa***
Logistic Model	0.2381	0.1463	0.1208	0.0339	0.0223	***0.11228***
Neural Network	0.1367	0.1069	0.2512	0.3280	0.4208	***0.24872***
SMO	0.2814	0.2819	0.2774	0.2473	0.1998	***0.25756***

Unsurprisingly, the same results as with distance analysis, average weighted Kappa results suggest that SMO is the better performing model amongst the three models, closely followed by neural network model. Logistic regression model still shows lowest performance in terms of agreement with Moody's.

5.7 Bubble Chart Analysis

Bubble charts were developed to facilitate interpretation of similarity in this research. Bubble charts enable a visualization of crosstabulation tables with clear localization of frequencies and a graphical representation of the observations through bubble size (see Hu and Ansell, 2005 for details). The bubble charts are presented in Figure 2.

Obviously, out of the three credit scoring approaches, neural network shows the weakest similarity to Moody's in 2004, since the bubble chart shows few bubbles are close to the diagonal line and most large size bubbles are away from the diagonal line. The logistic regression's bubble chart in 2000 can be interpreted in the same manner. In fact, the situation appears worse than the results for neural network model in 2004.

Conclusions from these bubble charts are confirmed by the distance and weighted Kappa results in sections 5.5 and 5.6. In 2004, the distance value from neural network approach is 1.881 (highest among three models) and weighted Kappa value is 0.1367 (lowest among three models). The situation is indeed worse for logistic regression model in 2000, since the distance value is 2.9167 and weighted Kappa value is 0.0223.

In addition, the similarity of research models to Moody's can also be analysed over time. The performance of the logistic regression model improves year by year from 2000 to 2004, as more large size bubbles are increasingly concentrated on the diagonal line. The opposite occurs for the neural network model in the same time period. Comparing with the trends of the other two credit scoring techniques, SMO shows a more consistent performance between 2000 and 2004.

Another interesting finding is that for all credit scoring techniques, the bubbles tend to move downwards year by year from 2000 to 2004. Bubbles above the diagonal line indicate higher ratings for Moody's than for research models. Bubbles below the diagonal line indicate lower rating for Moody's than for research models. Thus, adopting Moody's as a benchmark, it can be said that research models possibly underrate the credit situation of sample companies in 2000 and overrate the credit situation in 2004.

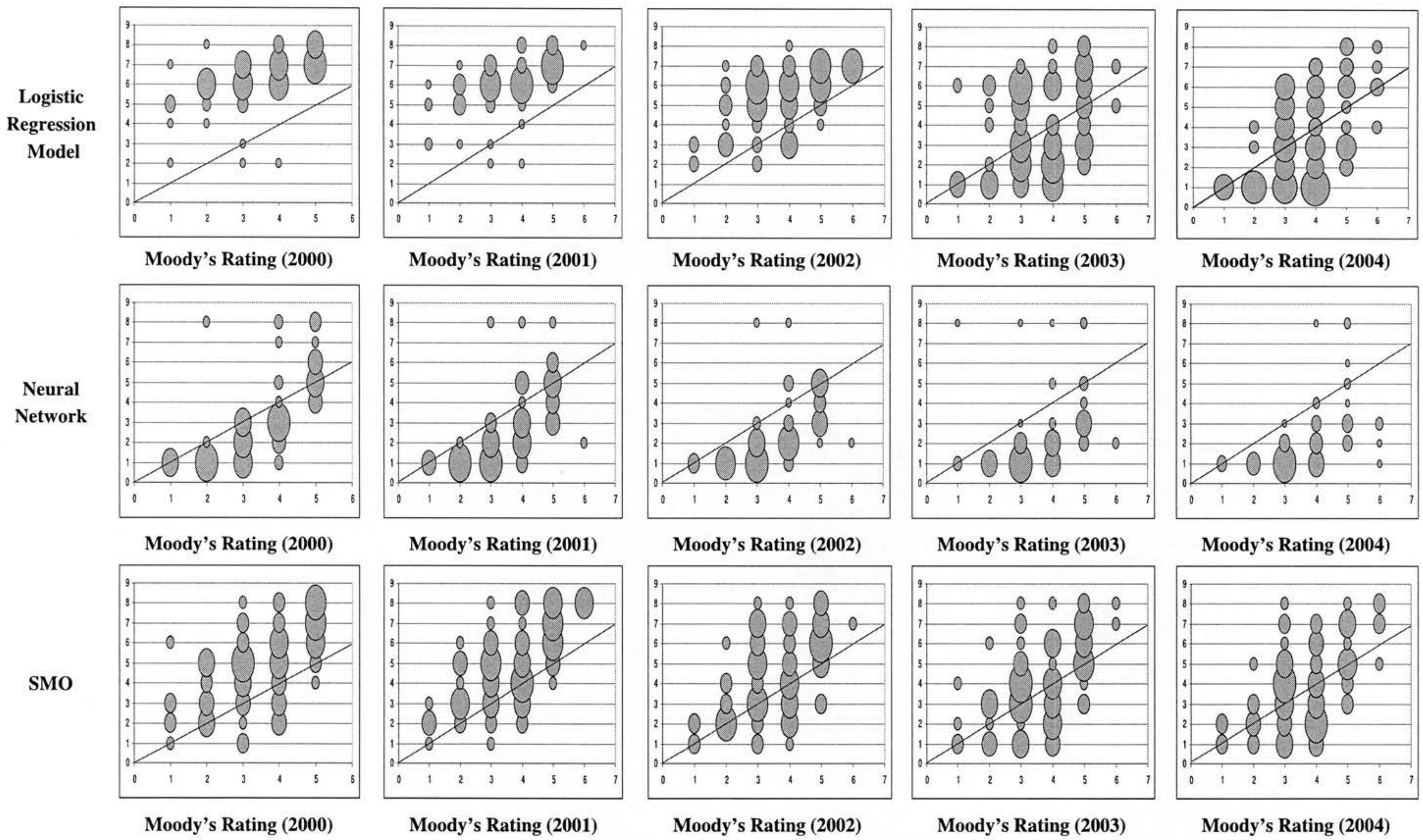


Figure 2 Bubble Charts Analysis

8. Discussions

This paper constructed a retail financial distress anticipatory model based on five key variables: Debt Ratio, Total Debt / (Total Debt + Market Capitalization), Total Assets, Operating Cash Flow and Government Debt / GDP, which proved to have sound classification performance in Hu and Ansell (2005).

US, European and Japanese markets are chosen for an international comparison analysis using five credit scoring methodologies, Naïve Bayes, Logistic Regression, Recursive Partitioning, Artificial Neural Network, and Sequential Minimal Optimization (SMO), over the time period from 2000 to 2004.

The international comparison analysis shows that regardless of the target countries, the average accuracy rates are above 86.5% and the average AUROC values are above 0.79. Moreover, model classification ability is only slightly different in the chosen countries. The results suggest that the five key variables have sound prediction ability in American, European and Asian retail markets.

When exploring the time dimension, all three market models possess best prediction ability in the year prior to financial distress with slight difference across markets. However, the longer the period before financial distress, the greater the difference across markets becomes, especially in terms of AUROC values.

The composite model was based on a dataset of 491 healthy and 68 distressed retail firms from USA, European and Japanese markets, over the time period from 2000 to 2004. Results show that all five credit-scoring techniques have the best classification ability in the year prior to the financial distress, with accuracy rates of above 88% and AUROC values of above 0.84. Furthermore, these techniques still remain sound five years before financial distress, as the accuracy rate is above 85% and AUROC value is above 0.72. However, it is difficult to conclude which modelling methodology has the absolute best classification ability, since the model's performance varies according to different time scales.

The findings above are potentially overly optimistic and may lead to overfitting, due to the limits of sample size. To overcome this problem, a series of comparison analysis using Moody's rating was performed. Based on the Kolmogorov-Smirnov significance test, distance measure, and weighted Kappa measure, it was found that SMO is the better performing model amongst the three models, closely followed by neural network model. Logistic regression model showed lowest performance in terms of similarity with Moody's. The bubble chart analysis also proved useful not only for comparing the similarity between two ordinal datasets, but also for detecting model performance trends. The results displayed consistent conclusions with other comparison techniques.

Thus far, the conclusions show a paradoxical result in that although the logistic model and the neural network model display better classification ability than the SMO composite model, the SMO composite model seems to be stronger in terms of comparability with Moody's rankings. A possible explanation is that the logistic regression model and neural network model fit the sample too closely, hence overfitting, whilst SMO does not.

In comparing the results from the international comparison analysis in this research with the findings in Hu and Ansell (2005), the performance of the USA model in this paper is similar to the model ability in Hu and Ansell (2005), despite different time periods. However, the performance of the European model and the Japanese model is worse than the model in Hu and Ansell (2005). A possible explanation is that as Hu and Ansell's (2005) model is based on the US market, USA model shows better performance than other market models. Moreover, the ability of the composite model is also worse than Hu and Ansell's (2005) model in terms of the AUROC value. This implies that a financial distress model has potentially better prediction ability when based on a single market. However, model construction is time-consuming and costly. Hence, global model development is still an important direction for future research. In this research, the composite model is only based on US, European and Japanese markets. More world retail markets can be included for future studies in order to ensure theoretical utility and practical applicability of the financial distress prediction models.

References

Altman, E.I. (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance*, 23, 4, 589-609

Altman, E.I. (1983) *Corporate Financial Distress: A Complete Guide to Predicting Avoiding and Dealing with Bankruptcy*, John Wiley and Sons Press, USA

Beaver, W.H. (1966) Financial ratios as predictors of failure, *Journal of Accounting Research*, 4, 3, 71-111

Betts, J. and Belhoul, D. (1987) The effectiveness of incorporating stability measures in company failure models, *Journal of Business, Finance and Accounting*, 14, 3, 323-334

Blum, M. (1974) Failing company discriminant analysis, *Journal of Accounting Research*, 12, 1, 1-25

Casey, C. and Bartczak, N. (1985) Using operating cash flow data to predict financial distress: some extensions, *Journal of Accounting Research*, 23, 1, 384-401

Coats, P.K. and Fant, L.F. (1993) Recognizing financial distress patterns using a neural network tool, *Financial Management*, 22, 3, 142-155

Cohen, J. (1960) A coefficient of agreement for nominal scales, *Educational and Psychological Measurement*, 20, 37-46.

Dawson, J. (2000) Retailing at century end: some challenges for management and research, *International Review of Retail, Distribution and Consumer Research*, 10, 2, 119-148

Deakin, E.B. (1972) A discriminant analysis of predictors of business failure, *Journal of Accounting Research*, 10, 1, 167-179

Deakin, E.B. (1976) Distributions of financial accounting ratios: some empirical evidence, *Accounting Review*, 51, 1, 90-96

Fan, A. and Palaniswami, M. (2000) Selecting bankruptcy predictors using a support vector machine approach, Paper presented at *2000 IEEE-INNS-ENNS International Joint Conference on Neural Networks*.

Fitch Ratings (2000) *Assigning Credit Ratings to European Retailers*, Fitch Ratings Press, USA (Downloadable from website: http://www.fitchratings.com/)

Fitch Ratings (2001) *Corporate: Corporate Rating Methodology*, Fitch Ratings Press, New York. (Downloadable from website http://www.fitchratings.com/)

Frydman, H., Altman, E.I. and Kao, D.L. (1985) Introducing recursive partitioning for financial classification: the case of financial distress, *Journal of Finance*, 40, 1, 269-291

Hand, D.J. (1997) *Construction and Assessment of Classification Rules*, John Wiley & Sons Ltd, Chichester, UK

Hamer, M. (1983) Failure prediction: Sensitivity of classification accuracy to alternative statistical methods and variable sets, *Journal of Accounting and Public Policy*, 2, 189-307

Hosmer, D.W. and Lemeshow, S. (2000) *Applied Logistic Regression*, Wiley Press, USA

Hu, Y.C., Ansell, J. (2005) Measuring retail company performance by using credit scoring techniques, submitted to *European Journal of Operational Research (EJOR)*. Paper presented at 2005 Credit Scoring and Credit Control IX, University of Edinburgh, 2005 Paris International Meeting (French Finance Association) and at 2006 WHU Campus for Finance, Germany

Hunt, S.D. (2000) *A General Theory of Competition: Resources, Competences, Productivity, Economic Growth*, Sage Publications Inc., USA

Mar-Molinero, C. and Serrano-Cinca, C. (2001) Bank failure: a multidimensional scaling approach, *European Journal of Finance*, 7, 2, 165-183

Marais, M.L., Patell, J.M. and Wolfson, M.A. (1984) The experimental design of classification models: An application of recursive partitioning and bootstrapping to commercial bank loan classifications, *Journal of Accounting Research*, 22 (Supplement), 87-114

McKee, T.E. (2003) Rough sets bankruptcy prediction models versus auditor signalling rates, *Journal of Forecasting*, 22, 8, 569-586

Mensah, Y.M. (1983) The differential bankruptcy predictive ability of specific price level adjustments: some empirical evidence, *Accounting Review*, 58, 2, 228-246

Mensah, Y.M. (1984) An examination of the stationarity of multivariate bankruptcy prediction models: A methodological study, *Journal of Accounting Research*, 22, 1, 380-395

Moody's Investors Service (1998) *Rating Methodology: Industrial Company Rating Methodology*, Moody's Investors Service Inc. Press, USA (Downloadable from website http://www.moodys.com/)

Moody's Investors Service (2002) *Retail Rating Methodology: Moody's Approach to Assessing Key Credit Issues in Retailing*, Moody's Investors Service Inc. Press, USA (Downloadable from website http://www.moodys.com/)

Moore, A.W. (2001) Cross-validation for detecting and preventing overfitting, Tutorial Slides, Carnegie Mellon University (Downloadable from website: http://www.cs.cmu.edu/)

Ohlson, J.A. (1980) Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research*, 18, 1, 109-131

Pantalone, C. and Platt, M. (1987) Predicting failure of savings and loan associations, *American Real Estate and Urban Economics Association Journal*, 15, 2, 46-64

Piesse, J. and Wood, D. (1992) Issues in assessing MDA models of corporate failure: a research note, *British Accounting Review*, 24, 33-42

Platt, J.C. (1999) *Fast Training of Support Vector Machines Using Sequential Minimal Optimization*, In Schölkopf, B., Burges, C.J.C., Smola, A.J. (1999) *Advances in Kernel Methods: Support Vector Machines*, MIT press, UK

Platt, H.D. and Platt, M.B. (1990) Development of a class of stable predictive variables, *Journal of Business Finance and Accounting*, 17, 1, 31-51

Ross, S.A., Westerfield, R.W. and Jaffe, J. (1999) *Corporate Finance*, Irwin/McGraw-Hill Press, USA

Standard and Poor's (2002) *Standard and Poor's 2002 Corporate Rating Criteria*, The McGraw-Hill Companies press, USA (Downloadable from website: http://www.standardandpoors.com/)

Standard and Poor's (2003) *Standard and Poor's 2003 Corporate Rating Criteria*, The McGraw-Hill Companies press, USA (Downloadable from website: http://www.standardandpoors.com/)

Standard and Poor's (2005) *Sovereign Risk Indicators: Glossary of Terms*, The McGraw-Hill Companies press, USA (Downloadable from website: http://www.standardandpoors.com/)

Tabachnick, B.G. and Fidel, L.S. (2000) *Using Multivariate Statistics*, Allyn and Bacon Press, UK

Taffler, R.J. (1982) Forecasting company failure in the UK using discriminant analysis and financial ratio data, *Journal of the Royal Statistical Society*, Series A, 145, 3, 342-358

Taffler, R.J. (1984) Empirical models for the monitoring of UK corporations, *Journal of Banking and Finance*, 8, 2, 199-227

Tam, K.Y. and Kiang, M.Y. (1992) Managerial applications of neural networks: the case of bank failure predictions, *Management Science*, 38, 7, 926-947

Thomas, L.C., Edelman, D.B. and Crook, J.N. (2002) *Credit Scoring and Its Applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Trigueiros, D. and Taffler, R. (1996) Neural networks and empirical research in accounting, *Accounting and Business Research*, 26, 4, 347-355

Zavgren, C.V. (1983) The prediction of corporate failure: The state of the art, *Journal of Accounting Literature*, 2, 1-37

Zhang, G.P., Hu, M.Y., Patuwo, B.E. and Indro, D.C. (1999) Artificial neural networks in bankruptcy prediction: general framework and cross-validation analysis, *European Journal of Operational Research*, 116, 1, 16-32